



Facultad de Ciencias Económicas y Empresariales
ICADE

UNA LOVEBRAND EN LA ERA DIGITAL: EL CASO TIKTOK

Autor: Sonsoles Martínez González
Director: Ana Isabel Jiménez Zarco

MADRID | Abril 2024

RESUMEN

Se conoce una *lovebrand* como aquella marca que logra establecer una conexión emocional fuerte y positiva con sus consumidores. Esta relación entre la marca y el cliente va más allá que una simple transacción, caracterizándose por una fuerte y profunda lealtad y apego emocional por parte de los clientes. Estas marcas son capaces de generar sentimientos de satisfacción, amor, felicidad y confianza, dando lugar a ventajas competitivas tanto a corto como a largo plazo, por lo que se trata de un objetivo para muchas marcas.

Por esta misma razón, en un mundo empresarial digital y altamente competitivo como el actual, este concepto ha adquirido una importancia crítica. Es por ello, que en este trabajo se estudia el concepto de *lovebrand* además de analizar el caso particular de *TikTok*, plataforma la cual ha logrado captar la atención de millones de usuarios en todo el mundo.

Palabras clave: *Lovebrand*, *TikTok*, lealtad, amor, sentimientos, emociones, ventaja competitiva.

ABSTRACT

A *lovebrand* is recognized as a brand that successfully establishes a strong and positive emotional connection with its consumers. This relationship between the brand and the customer goes beyond a simple transaction, characterized by a profound and unwavering loyalty and emotional attachment from the customers. These brands can evoke feelings of satisfaction, love, happiness, and trust, leading to both short-term and long-term competitive advantages, making it a coveted goal for many brands.

For this very reason, in today's digital and highly competitive business world, this concept has become critically important. Therefore, this work studies the concept of *lovebrand* in addition to analyzing the particular case of *TikTok*, a platform that has managed to capture the attention of millions of users worldwide.

Key Words: *Lovebrand*, *TikTok*, loyalty, love, feelings, emotions, competitive advantage.

ÍNDICE

1. INTRODUCCIÓN.....	1
1.1. CONTEXTUALIZACIÓN	1
1.2. OBJETIVOS GENERALES Y ESPECÍFICOS.....	1
1.2.1 Objetivos Generales.....	1
1.2.2 Objetivos Específicos.....	2
1.3. METODOLOGÍA	2
1.4. ESTRUCTURA DEL TRABAJO	3
2. MARCO TEÓRICO.....	4
2.1. CONCEPTO DE LOVEBRAND	4
2.1.1 El concepto de marca	4
2.1.2 El concepto de brand love.....	7
2.2. PERFIL DE UNA LOVEBRAND	11
2.2.1 Dimensiones del amor a la marca	11
2.2.2 Elementos para la creación de una lovebrand.....	13
2.2.3 Implicaciones de ser una Lovebrand.....	19
2.3 CÓMO CLASIFICAR UNA LOVEBRAND	21
3. ANÁLISIS EMPÍRICO.....	24
3.1. QUE ES TIKTOK Y CUAL ES SU HISTORIA.....	24
3.2. ANÁLISIS DE LA MARCA	25
3.3. METODOLOGÍA	28
3.2.1. Obtención	28
3.2.2. Text Mining.....	29
3.2.3. Sentiment Analysis.....	30
3.4. RESULTADOS DEL ANÁLISIS.....	31
3.3.1. Descripción de la base de datos obtenida.....	31
3.3.2. Text Mining.....	34
3.3.3. Sentiment Analysis.....	45
4. CONCLUSIONES.....	51
6. ANEXO.....	60

Índice de Ilustraciones

Ilustración 1: Categorías de relaciones consumidor-marca según fournier (1998) ..7	
Ilustración 2: Marco conceptual que comprende los elementos y expresiones de una lovebrand16	
Ilustración 3: Distribución de usuarios de TikTok en todo el mundo a enero de 2024, por edad y sexo25	
Ilustración 4: Redes sociales más populares a nivel mundial a enero de 2024, clasificadas por MAUs (en millones)27	
Ilustración 5: Imagen de las primeras diez observaciones de la muestra33	
Ilustración 6: Fechas de publicación de la reseñas33	
Ilustración 7: Imagen del corpus34	
Ilustración 8: Recuento de reseñas negativas y positivas34	
Ilustración 9: Tokens creados a partir de los primeros tres documentos35	
Ilustración 10: Imagen del dataframe con los dieciséis términos más frecuentes y su frecuencia36	
Ilustración 11: Gráfico de los cincuenta términos más frecuentes y sus respectivas frecuencias37	
Ilustración 12: Gráfico de nube de palabras38	
Ilustración 13: Gráfico de los dieciséis términos más frecuentes para reseñas positivas y negativas por separado39	
Ilustración 14: Key-word-in-context de las primeras dieciséis reseñas que incluyen “good”, “love” y “like” dentro del grupo de reseñas negativas40	
Ilustración 15: Key-word-in-context de las primeras dieciséis reseñas que incluyen “delete” en del grupo de reseñas positivas41	
Ilustración 16: Key-word-in-context de las primeras 16 reseñas que incluyen “banned” dentro del grupo de reseñas negativas43	
Ilustración 17: Los veinte bigramas más frecuentes dentro de las reseñas y sus respectivas frecuencias.....44	
Ilustración 18: Gráfico de frecuencias de bigramas para reseñas positivas y negativas por separado44	
Ilustración 19: Conteo de palabras positivas y negativas (lexicón Bing)46	
Ilustración 20: Puntuaciones de polaridad asignadas a cada palabra (lexicón AFINN).....48	

Ilustración 21: Porcentajes de las palabras asociadas a los ocho sentimientos (lexicón NRC)	50
--	-----------

Índice de tablas

Tabla 1: Tipos de relaciones consumidor-objeto.....	12
Tabla 2: Dimensiones principales del amor a la marca.....	13
Tabla 3: Análisis de las dimensiones presentes en los modelos de lovebrands con nivel de priorización.....	19
Tabla 4: Ítems de la escala de brand love	22
Tabla 5: Descripción de las principales variables	32
Tabla 6: Puntuaciones de polaridad asignadas a cada palabra (lexicón Bing)	46
Tabla 7: Puntuaciones de polaridad asignadas a cada palabra (lexicón Afinn)	47
Tabla 8: Puntuaciones de polaridad asignadas a cada palabra según la emoción (lexicón NRC)	49

1. INTRODUCCIÓN

1.1. CONTEXTUALIZACIÓN

Hoy en día, en un entorno empresarial altamente competitivo, las relaciones básicas que pueden existir entre cliente y marca han pasado a un siguiente plano. Es por ello, que las marcas buscan no solo ser reconocidas, sino amadas por sus clientes, lo que influye directamente en la fidelidad y el compromiso a largo plazo. Sin embargo, definir y medir estas relaciones resulta complicado debido a su naturaleza subjetiva y multidimensional. La comprensión de la complejidad de estas conexiones emocionales es crucial para desarrollar estrategias efectivas de *branding* y marketing permitiendo a las marcas diferenciarse y alcanzar sus objetivos.

A lo largo de los años se han realizado estudios a cerca del concepto de *lovebrand*, donde se han estudiado y analizado distintos factores cuantitativos y cualitativos como la identidad de marca o la creación de relaciones interpersonales sólidas con los clientes, con el objetivo de explicar cómo se puede lograr crear esa conexión emocional fuerte con la marca. Actualmente, nos encontramos en lo que conocemos como era digital, donde cada vez más, nuestro día a día se ve condicionado por las nuevas tecnologías, lo que exige a las marcas una mayor adaptación a las dinámicas cambiantes del mercado. En este contexto, el presente trabajo busca realizar un estudio a cerca del concepto *lovebrand* y sus características propias además de aportar un análisis adicional y actualizado, enfocado en el caso específico de TikTok, para contribuir a la comprensión más profunda de lo que constituye una *lovebrand* en la era digital. TikTok, sirve como un caso de estudio relevante para descifrar estrategias exitosas de *lovebranding* debido a que su éxito se basa en gran parte a su capacidad para entablar una relación con sus usuarios que va más allá que un simple servicio.

1.2. OBJETIVOS GENERALES Y ESPECÍFICOS

1.2.1 Objetivos Generales

El objetivo principal de este Trabajo de Fin de Grado es realizar un análisis exhaustivo para medir la percepción pública de la aplicación TikTok y así evaluar si esta se trata de una *lovebrand* o no. En primer lugar, se realizará un estudio teórico y profundo acerca del

término *lovebrand* con el objetivo de identificar sus principales dimensiones y atributos, además de las estrategias y claves que se pueden aplicar para su consecución.

En segundo lugar, mediante un análisis de información extraída de internet y la aplicación de técnicas de análisis de sentimiento a través de texto y minería de datos se trata de identificar aquellos factores determinantes que contribuyen a la construcción de TikTok como *lovebrand* en el ámbito digital. Este análisis se enfoca principalmente en evaluar si las distintas estrategias de personalización y el compromiso con los usuarios han permitido a TikTok consolidarse como *lovebrand* e identificar las principales emociones que la marca genera en sus usuarios.

1.2.2 Objetivos Específicos

- Estudiar el concepto de *lovebrand* e identificar sus principales características y atributos.
- Identificar las estrategias que aporten un mayor valor a la marca y proporcionar así, recomendaciones prácticas para la gestión empresarial.
- Identificar y analizar los distintos factores que contribuyen a la percepción pública de TikTok como *lovebrand*, a través de la aplicación de técnicas de análisis de texto, como el análisis de sentimiento basado en *machine learning* y análisis de frecuencia de palabras clave.

1.3. METODOLOGÍA

En cuanto a la metodología que se llevo a cabo, esta dividida en varias etapas que permiten alcanzar los objetivos establecidos en la investigación. En primer lugar, se establece un marco teórico respaldado por literatura académica y científica donde se contextualiza qué es el amor de marca y se definen las distintas características, requisitos, significado y beneficios de ser una *lovebrand*, así como las implicaciones de este estatus en un contexto empresarial y social. Este análisis teórico se centrará en aspectos cualitativos para comprender en profundidad este fenómeno.

En segundo lugar, se llevará a cabo un análisis empírico, donde se aplicarán una serie de técnicas analíticas y cuantitativas, para ello, se emplearán herramientas avanzadas de procesamiento de datos. A tal efecto se empleará una base de datos obtenida a través de la plataforma en línea *Kaggle*, donde los datos utilizados fueron recopilados directamente

de la plataforma de distribución digital de contenido, *Google Play*. Posteriormente, se aplicaron técnicas de minería de datos y análisis de texto, utilizando la herramienta RStudio, cuyo entorno requerirá la utilización de un lenguaje de programación R, con el objetivo de identificar las emociones más predominantes, evaluar el sentimiento general de los comentarios y detectar distintas tendencias en la opinión pública. Durante el procesamiento de datos y antes de llevar a cabo el análisis de sentimiento, para garantizar una mayor calidad en los datos se realizó una limpieza de los datos y reducción del ruido. Una vez realizada la limpieza, el análisis de sentimiento se basa en el procesamiento de aquellas palabras clave más repetidas en las publicaciones, lo que permite evaluar las razones principales por las que la marca TikTok se percibe, o no, como una *lovebrand*.

1.4. ESTRUCTURA DEL TRABAJO

El trabajo se divide principalmente en cuatro partes. En primer lugar, una breve introducción donde se abarca una contextualización, objetivos principales, metodología y finalmente la estructura. En segundo lugar, se estudiará en profundidad el término *lovebrand* estableciendo un sólido marco teórico. En tercer lugar, se llevará a cabo un análisis empírico enfocado en el caso de la marca TikTok. Finalmente, en la última parte del trabajo se desarrollarán las distintas conclusiones que se alcanzan tras la realización del estudio conceptual y análisis empírico realizado en los anteriores apartados del trabajo.

2. MARCO TEÓRICO

2.1. CONCEPTO DE LOVEBRAND

2.1.1 El concepto de marca

A lo largo de los años el concepto de ‘marca’ ha ido evolucionando con el tiempo, siendo considerado un término cada vez más amplio y multifacético, pasando de ser considerado un simple logo o eslogan, a una herramienta indispensable para alcanzar la diferenciación. Actualmente, nos encontramos en un entorno altamente competitivo y dinámico donde cada vez supone un mayor desafío empresarial conseguir diferenciarse de la competencia, por lo que la marca cobra un rol cada vez más importante dentro de la empresa.

Existen muchas definiciones en cuanto al concepto de marca, según la Ley Española de Marcas (2001), en el artículo 4 se define el concepto de marca como: *“todos los signos, especialmente las palabras, incluidos los nombres de personas, los dibujos, las letras, las cifras, los colores, la forma del producto o de su embalaje, o los sonidos, a condición de que tales signos sean apropiados para:*

- a) distinguir los productos o los servicios de una empresa de los de otras empresas y*
- b) ser representados en el Registro de Marcas de manera tal que permita a las autoridades competentes y al público en general determinar el objeto claro y preciso de la protección otorgada a su titular.”* (Ley 17/2001, de 7 de diciembre).

Así mismo, según se indica en una revisión de la literatura obra de Gallart et al. (2019) la American Marketing Association (1960) dio su primera definición del concepto, refiriéndose a la marca como: *“un nombre, símbolo o diseño, o una combinación de ellos, dirigida a identificar los productos o servicios de un vendedor o grupo de vendedores y a diferenciarlos del resto de competidores”*. Podemos comprobar que ambas definiciones citadas hasta el momento se refieren a la marca como logo y herramienta de diferenciación. Sin embargo, fueron muchos los autores como Maurya y Mishra (2012) que argumentaron que el concepto de marca se trata de algo más complejo y que no debería estar limitado a un símbolo, diseño u objeto visual, sino que debería incluir también otras características intangibles o no visuales. Es por ello por lo que, tras una serie de consideraciones la American Marketing Association (2016) elaboró una nueva definición de marca de manera que se consideraba una marca como: *“nombre, término, diseño, símbolo, o cualquier otra característica que identifique el bien o servicio de un*

vendedor y lo diferencie de los demás vendedores de la competencia". En esta nueva definición, se tienen en cuenta otras características de manera y no se está limitando la marca únicamente a aquellos aspectos visuales, sino que esta definición se expande para abarcar elementos intangibles que contribuyen a la percepción global de esta. Estos componentes intangibles que constituyen parte de la marca se han convertido en diferenciadores clave en el mercado actual, donde los consumidores tratan de buscar conexiones más significativas y profundas con las marcas que deciden apoyar. Esto significa que la administración de la marca debe tratar de constituir una estrategia integral que vaya más allá de la mera estética.

De la misma manera la marca hoy en día constituye uno de los activos con mayor impacto en la construcción de valor tanto a nivel productos como a nivel empresarial. Según un estudio realizado por Colmenares (2017) *"La agregación de valor a los productos y servicios constituye parte fundamental para el crecimiento y sostenimiento de la satisfacción de las necesidades de los clientes particulares cada día más exigentes"*, de la misma manera que *"la marca constituye un factor determinante, diferencial e inimitable para crear y agregar valor en los productos, servicios y empresa"* afectando de manera notable en el grado de influencia sobre el cliente. Un cliente mostrará una mayor predisposición a adquirir un determinado producto con el que sienta un cierto nivel de afinidad o apego emocional hacia la propia marca. Colmenares recalca que, por lo tanto, *"los clientes compran una estructura intangible: un pensamiento, un sentimiento, una sensación física e incluso, en algunos casos, una interacción social que se unen en una satisfacción placentera"*.

Cabe destacar a otros estudios como Toro (2007), quien dentro de esta misma línea de pensamiento se refiere a la marca como un activo intangible que pasa a convertirse en un activo estratégico de mayor valor para cualquier empresa y de importancia superlativa, que debe manejarse con proyección en función de los propósitos centrales de la propia empresa a largo plazo. De manera que, la marca debe estar siempre alineada con los valores y objetivos empresariales de forma que sea una representación fiel de la empresa en todos sus aspectos.

El valor asociado a una marca por otro lado es un concepto muy presente y puede ser influenciado por múltiples variables. Entre ellas destacan las inversiones en marketing y publicidad (Toro, 2007) y la calidad del producto (Del Río Lanza et al., 2002) como

factores esenciales para construir y sostener ese valor. La sinergia de estas inversiones y su duración a lo largo del tiempo son aspectos críticos para el valor de marca ya que pueden actuar como potencias fortalecedoras o debilitadoras del mismo.

Este valor de marca desemboca en lo que se conoce en el mundo empresarial como *brand equity* y que según lo dispuesto por Siabato & Olivia (2014) este concepto es percibido por el consumidor como el valor agregado positiva o negativamente a un producto o servicio mediante la asociación con su respectivo nombre comercial. No obstante, son muchos los autores como Toro (2007) y Del Río Lanza et al. (2002) que hacen hincapié en las dificultades a la hora de estimar el valor de marca y por lo tanto el *brand equity*.

Según los autores Del Río Lanza et al. (2002), cuyo estudio tiene el objetivo de contribuir al conocimiento del concepto de valor de marca a través de un marco teórico y análisis empírico, se identificaron tres principales perspectivas de análisis para la aproximación del valor de marca, estas son: (1) la perspectiva financiera, (2) la perspectiva centrada en el consumidor y (3) la perspectiva global. Estas perspectivas proporcionan una base sustancial para comprender en mayor medida el concepto de valor de marca. En cuanto a la perspectiva financiera, esta trata de cuantificar la aportación al valor económico de la empresa y se centra en la valoración monetaria de la marca, esta visión es un pilar fundamental a la hora de establecer decisiones estratégicas en el ámbito financiero. Por otro lado, la perspectiva centrada en el consumidor se enfoca más en las distintas perspectivas, actitudes y comportamientos que mantienen los mismos en relación con la marca, y cómo estos elementos intangibles incrementan el valor de la marca en el mercado. Por último, según lo establecido por los autores la perspectiva global incluye estas dos perspectivas y trata de comprender el valor de marca como un fenómeno integral que ofrece tanto ventajas competitivas como beneficios económicos a través de una serie de variables que los consumidores asocian con la marca.

Actualmente nos encontramos en un contexto social altamente dinámico, por lo que es esencial reconocer el valor de marca como un activo dinámico que evoluciona a medida que lo hacen las percepciones del mercado y tendencias de consumo. “*Con las marcas dinamizadas a su máxima expresión, las empresas evidenciarán su sostenimiento y su futuro*” (Toro, 2007). Esto se alinea con lo propuesto tanto por Del Río Lanza et al. (2002) que sugieren que las marcas deben centrarse en comportamientos de compra actuales,

como en lo que sugiere Toro al defender que la marca “*se debe manejar con proyección a partir de los propósitos centrales de la Empresa y a largo plazo*”.

2.1.2 El concepto de brand love

La satisfacción, conocido como el núcleo del concepto del marketing, ha sido el objetivo principal del marketing durante más de 50 años. Sin embargo, a lo largo de los últimos años y a raíz de varios descubrimientos, se ha llegado a la conclusión de que un cliente satisfecho puede no ser suficiente para el éxito continuo en el mercado altamente competitivo en el que nos encontramos en la actualidad (Carroll y Ahuvia, 2006). A medida que nos encontramos cada vez en un entorno con mayor oferta, surge la necesidad de establecer una relación con los consumidores que vaya más allá de un simple juicio cognitivo como es la satisfacción a la hora de comprar un producto o servicio. Es en este contexto donde nace la necesidad de generar *brand passion* (Swimberghe et al., 2014). Según lo expuesto por los autores Swimberghe et al., *brand passion* se considera la conexión emocional definitiva entre consumidor y marca, donde los consumidores llegan a involucrarse personalmente con la propia marca, a veces incluso más de lo que lo hacen con sus amigos o seres queridos.

Estudios como Fournier (1998) mencionan la existencia de relaciones entre consumidor y marca y proporciona marcos preliminares para la ejecución y gestión de la construcción de la teoría de la relación consumidor-marca. Fournier reconoce la naturaleza multifacética de la relación consumidor-marca y por ello, descompone las distintas relaciones entre consumidor y marca en seis categorías distintas:

ILUSTRACIÓN 1: CATEGORÍAS DE RELACIONES CONSUMIDOR-MARCA SEGÚN FOURNIER (1998)



Fuente: Fournier, 1998

Dentro de estas categorías se incluyeron la pasión y el amor como una de ellas y se estableció que existía una base afectiva que recordaba al amor en el ámbito interpersonal en el centro de todas las relaciones sólidas que se tuvieron en cuenta para el estudio. Además, Fournier hace hincapié en que ambos sentimientos de amor y pasión son más profundos y duraderos que la simple preferencia, en lo que respecta al efecto que estos generan en el tipo de relación consumidor - marca.

Aquí es cuando aparece lo que conocemos como ‘*brand love*’ o ‘amor a la marca’ y que Carroll y Ahuvia (2006) definen como “*el grado de apego emocional apasionado que un consumidor satisfecho tiene por un nombre comercial en particular*”. De esta misma manera, los autores Carroll y Ahuvia (2006) recalcan que existen una serie de características que caracterizan el amor por la marca y que son constituidas por (1) la pasión por la marca, (2) el apego a la marca, (3) la evaluación positiva de la marca, (4) las emociones positivas en respuesta a la marca y (5) las declaraciones de amor hacia la marca.

De igual manera, es importante recalcar las diferencias que existen entre una relación interpersonal y una relación entre consumidor y marca ya que, el brand love es un concepto que involucra una serie de sentimientos unidireccionales del consumidor hacia la marca, a diferencia de una relación interpersonal que se caracteriza por ser bidireccional (Bergkvist y Bech-Larsen, 2010). Esto no quiere decir que una persona no pueda llegar a alcanzar un sentimiento tan profundo como es el sentir amor hacia una marca.

Dentro de este marco conceptual cabe destacar a otros estudiosos como Sallam y Wahid (2015), que consideran el brand love como “*un constructo de marketing reciente, que se ha demostrado que influye en variables de marketing importantes como el boca a boca y la toma de decisiones de compra*”, para demostrarlo ambos autores realizaron un estudio con el objetivo de explorar el efecto de lo que denominaron como variables antecedentes al brand love y entre los resultados principales se identificó la lealtad a la marca como un mediador parcial entre el brand love y lo que conocemos como valor de marca, en otras palabras, cuando los consumidores desarrollan amor por la marca, esto conduce parcialmente a una mayor lealtad hacia esa marca, y esta lealtad, a su vez, aumenta el valor de marca. Sin embargo, los autores no dejan de reconocer la importancia de la satisfacción como variable antecedente al sentimiento de amor y establecen que “*la*

satisfacción del cliente y la identificación con la marca desempeñan un papel importante en la construcción del amor del consumidor hacia la marca” (Sallam y Wahid, 2015).

Por otro lado, Carroll y Ahuvia (2006) hacen hincapié en la importancia que reside en poder distinguir entre los términos de satisfacción y amor a la marca, el primero siendo un constructo más macro, lo que quiere decir que no todos los consumidores satisfechos sienten amor hacia la marca y además establecen cuatro diferencias clave entre ambos términos. En primer lugar, catalogan la satisfacción como un simple juicio cognitivo mientras que se dice que el amor por la marca consta de un enfoque afectivo mucho más fuerte. En segundo lugar, se refieren a la satisfacción como un mero resultado fruto de la transacción, mientras que el amor a la marca es a menudo el resultado de una relación a largo plazo entre el consumidor y la marca. En tercer lugar, se menciona que la satisfacción está vinculada en muchas ocasiones a la confirmación de expectativas mientras que el amor a la marca no requiere ningún tipo de expectativa ni desconfirmación. Por último, los autores establecen que a diferencia de los requisitos que envuelven al término satisfacción, el amor por la marca involucra la voluntad de declarar amor e implica del mismo modo la integración de la marca en la identidad del propio consumidor.

Dentro de estas mismas líneas de pensamiento los autores Coelho et al. (2019) recalcan que el amor a la marca se caracteriza por tratarse de algo íntimo y emocionante marcado por un alto nivel de compromiso y una serie de esfuerzos significativos por parte del consumidor para obtener posesión de dicha marca. Tal devoción puede incluso llegar a bordear lo que se conoce como fanatismo, esto en muchas ocasiones lleva a los consumidores a rechazar otras alternativas y a mantenerse fieles a la que consideran como su marca de confianza (Coelho et al., 2019). Los autores Coelho et al. (2019) identificaron también como a través de las comunidades de marca, donde los distintos miembros interactúan entre ellos de manera diaria, se consigue integrar a los individuos con la marca de manera que estos se sientan parte de ella y donde sus participantes se resisten a terminar una relación, de la misma manera que cuando un individuo ama una marca y crea una conexión profunda con ella, es muy complicado que este rompa su vínculo. De esta manera, el valor asociado a la marca se desarrolla y crece progresivamente dentro de la comunidad.

Al adentrarnos más en lo que respecta a la relación amorosa entre consumidor y marca, nos topamos con el término de lovebrand. En cuanto a la diferencia entre brand love y lovebrand, los autores Lindenmeyer et al. (2023) recalcan que mientras que el brand love describe el sentimiento que los consumidores tienen hacia la marca en cuestión, una lovebrand se trata de hacia dónde se dirige este sentimiento. Este término está estrechamente vinculado a la teoría existente sobre el amor a la marca y está establecido en algunos sectores corporativos, pero es relativamente nuevo en lo que respecta a la literatura académica (Lindenmeyer et al., 2023), por lo que actualmente no existe un conocimiento profundo acerca del mismo en el aspecto académico.

Existen varias definiciones para el término lovebrand, por ejemplo, autores como Lindenmeyer et al. (2023) se han referido al término como *“un concepto subjetivo al cliente individual que siente una conexión profunda, apasionada y emocional hacia la marca, la cual se alinea con su propia personalidad e integra la marca en su propia identidad”*. Por otro lado, Aguado (2016) define el concepto de lovebrand como *“aquellas marcas que han conseguido atraparnos, bien por su imagen, por lo que hacen por nosotros, por el beneficio que aportan (consciente o inconsciente), por cómo nos hacen sentir, y no sólo las consumimos, sino que damos un paso más allá y nos hacemos Brand Ambassadors voluntarios por convicción”*. Aguado (2016) recalca también la diferencia entre las lovebrand y lo que denominó como las *invisible brands* *“marcas que ni vemos”* o las *everyday brands* *“marcas cercanas, que han entrado en nuestra vida hace tiempo y que tenemos integradas en nuestro día a día”*.

Ambas definiciones del término lovebrand citadas demuestran que en este contexto se establece una relación que va más allá del nivel funcional, dando lugar a una relación más intensa en términos tanto personales como emocionales con el consumidor. No se trata de una marca que cumple con las expectativas en cuanto a calidad y precio, sino que se trata de aquella que consigue desarrollar una determinada narrativa y logra resonar con las historias personales de los consumidores. Es por ello por lo que para los especialistas en marketing la creación de una lovebrand se considera tanto una oportunidad como un desafío en muchos ámbitos.

2.2. PERFIL DE UNA LOVEBRAND

2.2.1 Dimensiones del amor a la marca

Según la Teoría Triangular del Amor de Sternberg que fue desarrollada en 1986, se puede dividir el amor en tres dimensiones:

- i. **Intimidad** “abarca los sentimientos de cercanía, conexión y vinculación que uno experimenta en las relaciones amorosas” (Sternberg, 1986),
- ii. **Pasión** “abarca los impulsos que conducen al romance, la atracción física, y consumación sexual” y
- iii. **Compromiso** “abarca, en el corto plazo, la decisión de amar a otro, y en el largo plazo, el compromiso de mantener ese amor”.

Más adelante Sternberg dictamina que estas tres dimensiones desencadenan ocho posibles tipos de amor (No-amor, gusto, amor enamorado, amor vacío, amor romántico, amor acompañado, amor fastuoso y amor consumado) en consecuencia de la presencia o ausencia de los tres componentes citados.

Tomando esta teoría como precedente y fuente de inspiración, los autores Shimp y Madden (1988) proponen un modelo conceptual sobre las relaciones consumidor-objeto donde las tres dimensiones establecidas por Sternberg pasan a convertirse en (i) agrado, (ii) anhelo y (iii) decisión/compromiso en un contexto de consumo. Se establece que cuando se dan estos tres componentes, estos contribuyen fuertemente a la lealtad hacia el objeto, en este caso la marca (Albert et al, 2008).

El componente de agrado se refiere a emociones tales como el afecto y cariño, además de incluir aquellos aspectos establecidos por Sternberg asociados a la intimidad. El anhelo por otro lado se refiere a tener un deseo serio o fuerte por algo representando un continuo de posibilidades. Finalmente, el componente de decisión/compromiso se trata de un aspecto más cognitivo, es el que menos varía en cuanto a la definición proporcionada por Sternberg, esto quiere decir que en el corto plazo un consumidor puede decantarse por una marca porque sus atributos o beneficios son justo lo que estaban buscando y más adelante en el largo plazo comprometerse con la marca como una cuestión de lealtad (Shimp y Madden, 1988).

Shimp y Madden (1988) identifican los siguientes tipos de relaciones de consumo, consecuencia de la presencia o ausencia de los tres componentes actualizados:

TABLA 1: TIPOS DE RELACIONES CONSUMIDOR-OBJETO

Tipo de relación	Componente		
	Agrado	Anhelo	Decisión/Compromiso
Desagrado	-	-	-
Gusto	+	-	-
Enamoramamiento	-	+	-
Funcionalismo	-	-	+
Deseo inhibido	+	+	-
Utilitarismo	+	-	+
Deseo cedido	-	+	+
Lealtad	+	+	+

Fuente: Shimp y Madden (1988)

En la tabla podemos observar la evolución de las distintas relaciones de consumo empezando por ‘desagrado’ donde no se aplican ninguno de los tres componentes, según Shimp y Madden esto da a entender que el consumidor no muestra ningún tipo de sentimiento en particular hacia la marca. Por último, se observa el tipo de relación denominado ‘lealtad’ donde confluyen todas las dimensiones y donde los autores especifican que el consumidor demuestra una relación íntima con la marca, un fuerte deseo por comprarla y se compromete (por lo menos a corto plazo) a apoyar a la marca en cuestión, esta relación se corresponde por lo tanto con el objetivo al que cualquier marca aspira.

Para comprender más a fondo el concepto de amor a la marca hace falta adentrarse en la psicología del consumidor, permitiéndonos comprender mejor los distintos tipos de relaciones que estos mantienen con las marcas que aman. En este sentido fueron pioneros los autores Albert et al. (2008) que en su obra ofrecen una visión exhaustiva de las distintas dimensiones que componen este amor hacia la marca. Como resultado del

análisis se identifican 11 dimensiones principales del amor hacia la marca que se pueden ver a continuación:

TABLA 2: DIMENSIONES PRINCIPALES DEL AMOR A LA MARCA

Dimensiones	
Pasión (por la marca)	Atracción (sentimiento hacia la marca)
Duración de la relación (la relación con la marca existe desde hace mucho tiempo).	Singularidad (de la marca y/o de la relación)
Auto congruencia (congruencia entre la autoimagen y la imagen del producto)	Belleza (de la marca)
Sueños (la marca favorece los sueños del consumidor)	Confianza (la marca nunca ha defraudado)
Recuerdos (evocados por la marca)	Declaración de afecto (sentimiento hacia la marca)
Placer (que la marca proporciona al consumidor)	

Fuente: Albert et al. (2008)

Sin embargo, Albert et al. (2008) establecen que, con toda probabilidad, para que exista una relación amorosa entre consumidor y marca no hace falta que estén presentes simultáneamente todas las dimensiones.

2.2.2 Elementos para la creación de una lovebrand

En esencia, una lovebrand, en términos básicos, se trata de una marca capaz de suscitar un sentimiento que vaya más allá de lo racional y tan profundo como es el amor e inspiración en sus consumidores. Muchas marcas aspiran a alcanzar esta posición, no obstante, actualmente se carece de una fórmula que indique de manera explícita los pasos a seguir para conseguir alcanzar dicho estatus. Sin embargo, existen teorías y estudios como los de Roberts (2004), Aguado (2016) o Lindenmeyer et al. (2023), que aportan una serie de descubrimientos e indicaciones que pueden servir de guía a las marcas en caso de que estas quieran convertirse en lovebrands.

Según Roberts (2004), existen tres elementos clave que convierten a una marca en una lovebrand:

1. **Misterio:** En muchas ocasiones las empresas tratan de mitigar el misterio con el objetivo de aumentar su credibilidad en el mercado, sin embargo, el autor se refiere al misterio como una oportunidad, además de un elemento clave para convertirse en una lovebrand a través de lo que denominó “los cinco magníficos del misterio”:
 - a. **Grandes historias:** El autor ve el misterio como una herramienta persuasiva que permite a la marca captar la atención del consumidor, adaptándose a las nuevas generaciones, pero manteniéndose fiel a su historia.
 - b. **Pasado, presente y futuro:** *“las lovemarks son como las grandes familias, combinan las lecciones del pasado y la dinámica del presente para crear un gran futuro”* (Roberts 2004). El autor menciona lo importante que es para una marca su legado emocional y utilizarlo a su favor, ya que este está estrechamente vinculado a las relaciones que mantienen con sus consumidores.
 - c. **Despierta los sueños:** Según Roberts, una marca que conoce los sueños de sus consumidores es porque estos confían en ella y la aman. Además, *“despertar los sueños es una poderosa manera de mostrarle a las personas que entendemos sus deseos y que podemos hacerlos realidad”* (Roberts, 2004).
 - d. **Mitos e iconos:** El autor califica los mitos e iconos como símbolos memorables y directos que pueden servir como un puente directo hacia el corazón de los consumidores.
 - e. **Inspiración:** Se dice que la inspiración está directamente vinculada con el amor y que puede transformar vidas. Para demostrarlo se hace alusión a los equipos deportivos locales que todos hemos ido a ver alguna vez y a las emociones que estos generan en sus fieles seguidores.
2. **Sensualidad:** En cuanto a este elemento, según Roberts, los sentidos nos recuerdan quienes somos y constituyen la vía más eficaz para acceder a un corazón, además son lo que nutre nuestro conocimiento y es muy difícil de anular. Sin embargo, en muchas ocasiones las empresas tratan de aumentar el aspecto sensorial a través de vías rápidas como aumentando el sonido, el brillo, el sabor o la textura, esto hace que los consumidores desconecten y se pierdan por el camino.

Para evitar estas situaciones debe fomentarse la sensualidad desde una perspectiva más humana. Para lograr esa conexión tan ansiada con el consumidor, el autor indica que hace falta estimular varios sentidos al unísono, y es que cuando los sentidos trabajan de la mano, es cuando aparece nuestro sexto sentido, es decir, la intuición.

3. **Intimidad:** La cercanía y confianza son aspectos que buscamos en nuestras relaciones y que están directamente relacionado con la intimidad. El autor menciona cómo en numerosas ocasiones, muchas marcas han perdido sus vínculos emocionales al desaprovechar oportunidades de acercamiento íntimo con aquellos consumidores que las amaban para poner el foco en otro público objetivo: los accionistas. Este es un aspecto más polémico y escurridizo que los anteriores dos al tratarse de un proceso de doble sentido, hay que saber escuchar y hablar, la mayoría de las marcas solo practican lo segundo. *“La intimidad exige una profunda comprensión de lo que le importa a la gente. Lo que supone que también nosotros debemos estar dispuestos a revelar quienes somos y confesar nuestros propios sentimientos”* (Roberts, 2004). No podemos esperar que los consumidores se abran con la marca sin recibir nada a cambio. Tal y como el autor establece, la intimidad consta de tres facetas diferentes:

- a. **Empatía:** *“Para comprender y responder a las emociones de los demás”* (Roberts, 2004),
- b. **Compromiso:** *“Prueba que estamos en una relación a largo plazo”* y
- c. **Pasión:** *“Esa chispa luminosa que mantiene viva la relación”*.

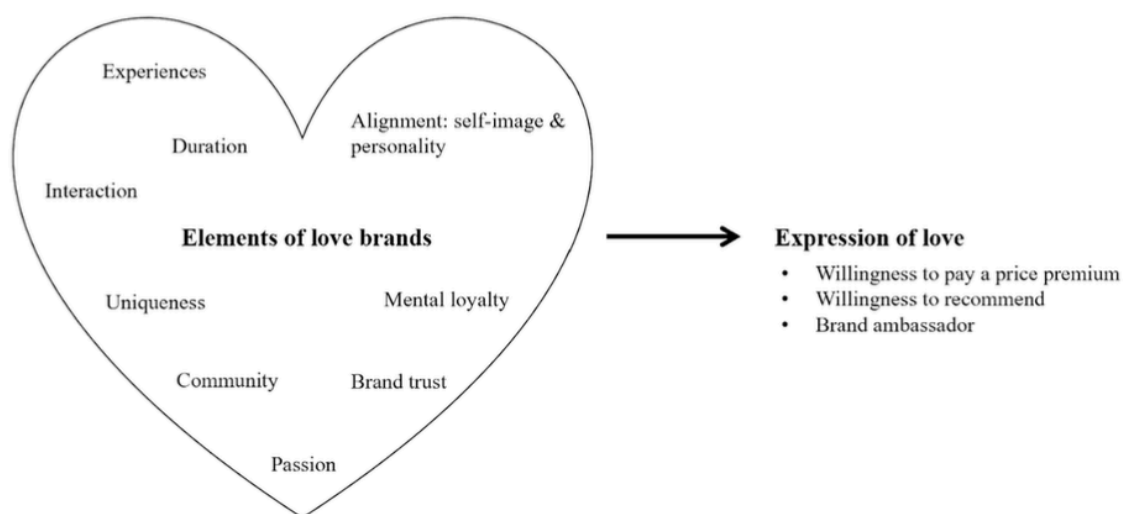
Una vez analizados los tres elementos principales que caracterizan a una lovebrand según Roberts, no se debe pasar por alto el factor de respeto, que también es desarrollado por el autor más adelante. El autor llegó a la conclusión de que donde no haya respeto nunca se desarrollará una relación verdadera y para que exista una lovebrand deben coexistir al unísono tanto el respeto como el amor hacia la marca.

Unos años más tarde Aguado (2016) sugirió varias claves para aquellas marcas que quisieran convertirse en lovebrands. En primer lugar, se menciona la aplicación de estrategias *bottom-up*, estas son estrategias que parten del consumidor en lugar del producto. Es decir, se trata de darle al consumidor lo que quiere para conquistarle y no

imponerlo sobre él de manera que el consumidor cuenta con un mayor poder y es libre de tomar las decisiones oportunas según su propio criterio. Más adelante se recalca la importancia de crear experiencias para el cliente como medio para establecer una conexión profunda y la relevancia actual de lo que se conocen como *experience brands*, siendo estas marcas las que “ponen tanta atención en la experiencia alrededor del producto o servicio, como en las características del mismo. Marcas que se conocen más por lo que hacen, que por lo que son. Marcas que te hacen sentir” (Aguado, 2016). En la sociedad tan dinámica y cambiante en la que nos encontramos actualmente, poner el foco en el consumidor y la creación de experiencias y sensaciones para el cliente son dos aspectos clave a la hora de crear una conexión fuerte y duradera, y pueden constituir iniciativas sólidas que permitan establecer las bases para la creación de una lovebrand.

Esta cuestión también es abordada por Lindenmeyer et al. (2023), quienes con el objetivo de identificar patrones de conducta en las relaciones consumidor-marca e identificar aquellos denominadores comunes que contienen las lovebrand, identificaron (tras una serie de entrevistas) nueve elementos fundamentales que constituyen una lovebrand. Adicionalmente, los autores enriquecieron su análisis a través de la inclusión de una serie de expresiones vinculadas al amor que un consumidor muestra por la marca, en caso de que se cumpliesen todas las condiciones. Estos elementos y expresiones de amor se pueden ver a continuación en el siguiente marco conceptual:

ILUSTRACIÓN 2: MARCO CONCEPTUAL QUE COMPRENDE LOS ELEMENTOS Y EXPRESIONES DE UNA LOVEBRAND



Fuente: Lindenmeyer et al., 2023

Este marco sirve como guía comprehensiva para aquellas empresas que buscan cultivar o nutrir una lovebrand, enriqueciendo el concepto y obteniendo resultados positivos. Cabe mencionar que en muchas ocasiones estos elementos son iguales o se asemejan a las dimensiones del amor a la marca mencionadas anteriormente y desarrolladas por los autores Albert et al. (2008) y que demuestran una vez más lo vinculados que están el concepto de lovebrand con la teoría existente del amor a la marca a pesar de tratarse de conceptos diferentes como se indica de la misma manera en apartados anteriores.

Además, se debe tener en cuenta la diferencia de perspectivas en los artículos académicos frente a los corporativos. Mientras que los estudios corporativos como el de Roberts (2004), se centran en aspectos más generales como el misterio, la sensualidad y la intimidad que pueden emplearse para describir el amor en general, la perspectiva académica aborda el concepto desde una perspectiva más amplia de las emociones, sin centrarse en detalles específicos (Lindenmeyer et al., 2023).

En cuanto a los elementos identificados por Lindenmeyer et al. (2023):

- a) **Comunidad:** Participar en una comunidad de aficionados fortalece el sentimiento de pertenencia entre los consumidores miembros, que comparten una conexión especial con esa marca.
- b) **Experiencias:** En línea con lo expuesto por Aguado (2016), los autores Lindenmeyer et al. consideran que las experiencias vividas con las marcas juegan un papel crucial a la hora de desarrollar una conexión fuerte con el cliente, similar a cómo las relaciones personales se ven fortalecidas por experiencias compartidas.
- c) **Alineamiento (autoimagen y personalidad):** Los consumidores tienden a vincular la propia marca con atributos deseados para su propia imagen. De manera que el estudio muestra que una lovebrand puede llegar a integrarse en la personalidad del individuo de manera que la marca acaba alineándose con su propia imagen.
- d) **Duración:** De la misma manera en que una relación interpersonal se fortalece y forja a medida que pasan los años, la conexión y los vínculos hacia una marca crecen gradualmente. Los autores recalcan que el amor hacia una marca no implica necesariamente la adquisición de sus productos y que una marca con historia y reputación fuerte puede transformarse en una lovebrand con el tiempo.

- e) **Pasión:** Para incentivar el enamoramiento de los consumidores es de vital importancia fomentar la pasión hacia la marca, de manera que proporcionando exclusividad al consumidor en la relación se puede agilizar el desarrollo de una lovebrand.
- f) **Singularidad:** Según mencionan los autores, en muchas ocasiones los entrevistados a menudo estaban dispuestos a pagar un precio *premium* por los productos en caso de no encontrar otros equivalentes, así como la predisposición a recomendar, convirtiéndose en embajadores de la marca dentro de su comunidad.
- g) **Lealtad mental e Interacción:** Las interacciones regulares fomentan la lealtad a la marca. Esto se debe a que el amor generado hacia la marca cuenta con un aspecto psicológico y no está vinculado a la frecuencia de compra, sino que a los puntos de contacto que se tiene con la marca en el día a día.
- h) **Confianza:** Este supone el elemento final dentro del marco propuesto por Lindenmeyer et al. debido a que los consumidores se enorgullecen de sus lovebrands designadas, confían en ellas y desean contribuir en su desarrollo.

Cabe mencionar que las entrevistas se realizaron sobre sujetos y marcas que muestran diferencias entre sí y dependiendo del contexto de la persona misma, su estilo de vida u otros factores externos se puede generar amor a la marca o no. Esto resalta la importancia de la subjetividad dentro del concepto de lovebrands (Lindenmeyer et al., 2023).

Antes de dar por finalizado este análisis de los distintos elementos característicos de una lovebrand, es importante hacer mención del estudio realizado por Montalvo-Arroyave (2023) donde se recoge una perspectiva sólida respecto a la estructura teórica que da base a este concepto. Este estudio permite obtener una mejor comprensión de lo expuesto hasta día de hoy acerca del concepto de lovebrand y aportar al desarrollo futuro de estas marcas. Para ello, se analizan las dimensiones presentadas en 54 modelos diferentes y que abordan las dimensiones para la creación de lovebrands. Cabe mencionar que para elaborar esta perspectiva se incluyeron estudios que evalúan las distintas dimensiones que caracterizan el amor a la marca debido a lo estrechamente vinculados que están ambos conceptos (al ser una lovebrand consecuencia del amor hacia la marca).

Los resultados del análisis muestran las siguientes 10 dimensiones como las más presentes en los modelos, categorizadas como antecedentes, es decir, elementos que

puedan ser catalizadores de una relación futura con consumidores que desemboque en una lovebrand:

TABLA 3: ANÁLISIS DE LAS DIMENSIONES PRESENTES EN LOS MODELOS DE LOVEBRANDS CON NIVEL DE PRIORIZACIÓN

Dimensiones	Modelos que han utilizado cada dimensión (Uds.)	Uso de cada dimensión en los modelos (%)
Identidad de/con la marca	24	44%
Experiencia de marca	15	28%
Pasión	11	20%
Intimidad	11	20%
Imagen de marca	9	17%
Autenticidad de marca	8	15%
Calidad	6	11%
Marca hedónica	6	11%
Sentido de comunidad	6	11%
Sensualidad	6	11%

Fuente: Montalvo-Arroyave, 2023

En la Tabla 3 se aprecia que, dentro de las dimensiones citadas con mayor frecuencia en los estudios analizados acerca de esta literatura, en numerosas ocasiones se incluyen los expuestos por los autores que se mencionan a lo largo de este segmento. La investigación desarrollada por Montalvo-Arroyave nos permite obtener una visión más amplia de los diversos componentes que definen a una lovebrand, basándose empíricamente en la literatura existente y sin juicio propio, a diferencia de Lindenmeyer et al., o Roberts en sus respectivas obras.

2.2.3 Implicaciones de ser una Lovebrand

Una vez se alcanza la posición de lovebrand, surgen una serie de consecuencias que se exponen a continuación. Para ello, Sarkar (2011) hace referencia en su estudio a cuatro consecuencias principales fruto del amor a la marca.

En primer lugar, y en línea por lo expuesto por Lindenmeyer et al. (2023), se dice que, como consecuencia del apego emocional del consumidor hacia la marca, este estará **dispuesto a pagar un precio *premium*¹** por el producto o servicio el cual variará en función del grado de amor que este sienta por la marca.

En segundo lugar, Sarkar (2011) establece que el amor por la marca influye positivamente en la **lealtad hacia la marca**. Se menciona que “*el amor por la marca predice la lealtad hacia la marca*” (Carroll y Ahuvia, 2006) y que cuanto mayor sea este amor hacia la marca, más fuerte será esta lealtad y más barreras estará dispuesto a superar el consumidor para adquirir el producto o servicio.

El autor establece como tercera consecuencia el ***Word-of-mouth (WOM)*² positivo**, esto fue demostrado por Carroll y Ahuvia (2006). El *WOM* acelera la difusión de la marca en el mercado, se asume que “*si el consumidor ama una marca, entonces también se espera que el consumidor hable positivamente de la marca a otros*” (Sarkar, 2011).

Por último, como última consecuencia de las lovebrands, Sarkar (2011) une dos de los beneficios previamente mencionados y hace referencia a la **vinculación entre la lealtad de acción y el *WOM***. Esto se debe a que un consumidor leal se sobreentiende que está satisfecho con la marca y eso motivará al consumidor a difundir *word-of-mouth* positivo (Sarkar, 2011).

Así mismo, es importante hacer referencia a la digitalización y en cómo las consecuencias de ser una lovebrand se han visto afectadas por ella. Se dice que las redes sociales han dado lugar a una nueva manera de construir relaciones entre las marcas y sus consumidores y que pueden conectar a ambos, fomentando así el compromiso del consumidor (Hollebeek et al., 2014; Salem et al., 2019). Debido a esto, muchas marcas invierten grandes cantidades de dinero en la imagen que transmiten a sus consumidores a través de la pantalla. Como indican Hollebeek et al., las redes sociales pueden fomentar el compromiso del consumidor con la marca, lo que obras como Coelho et al. (2019),

¹ Cuando un consumidor está dispuesto a pagar un precio premium por un producto o servicio significa que está dispuesto a pagar más dinero por este producto/servicio del que normalmente gastaría en otros productos o servicios similares de otras marcas.

² *Word-of-mouth*, conocido en español como el boca a boca.

Sternberg (1986) o Roberts (2014) establecen como parte fundamental dentro de las relaciones de amor.

Además, el internet y las redes sociales constituyen hoy en día las principales fuentes del boca a boca en la actualidad, esto está directamente relacionado con las consecuencias de ser una lovebrand como se indicó previamente. En el panorama actual, cualquiera puede acceder a internet y publicar su opinión acerca de una marca y que lo vea cualquier otro usuario sin necesidad de conocerse entre ambos. He aquí cuando las marcas ponen el foco en los *influencers*³ y a la enorme influencia que estos generan en su público. Según Jiménez-Castillo & Sánchez-Fernández (2019) los *influencers* influyen en las percepciones, valoraciones e intenciones de compra de sus seguidores (en sus respectivos roles como consumidores) sobre las marcas que este recomienda. Es por ello por lo que para que una marca alcance el papel de lovebrand está tendrá que saber aprovechar y adaptarse a estos cambios.

2.3 CÓMO CLASIFICAR UNA LOVEBRAND

A lo largo de este estudio hemos analizado numerosas perspectivas de distintos autores en cuanto a que es una lovebrand y en qué factores debemos enfocarnos en caso de querer alcanzar este estatus. Sin embargo, en nuestro caso, para medir este concepto de manera efectiva emplearemos el estudio “*Brand Love: Development and Validation of a Practical Scale*” obra de Bagozzi et al. (2017), como referencia para aproximar hasta qué punto una marca (en este caso TikTok), alcanza la posición de lovebrand. Este estudio en particular es muy útil, ya que desarrolla y valida una escala de amor hacia la marca capaz de predecir la lealtad, el *word-of-mouth* y la resistencia a la información negativa por parte de los consumidores y que podemos aplicar sobre diversas marcas y contextos. (Bagozzi et al. 2017). Esto último nos resulta especialmente útil debido a que el objeto de estudio en el trabajo en este caso se trata de una plataforma digital, más concretamente una red social.

Se trata de una escala más parsimoniosa que la de Batra et al. (2012) (obra citada en numerosos estudios), debido a que para elaborar esta escala se realizaron varias pruebas de validez importantes no consideradas por Batra et al. Además, la escala desarrollada por los autores Bagozzi et al. cuenta con tres versiones anidadas de 26, 13 y 6 ítems a

³ Un *influencer* es aquel que posee el poder de influencia sobre otros usuarios debido a que estos lo ven como referente.

diferencia de la de Batra et al. (2012) que cuenta con una sola versión constituida por 56 ítems, medida la cual es simplemente demasiado larga para ser pragmáticamente útil en muchos estudios (Bagozzi et al. 2017).

A continuación, en la tabla 4 se muestran los distintos factores e ítems de las versiones anidadas de 13 y 6 componentes elaborados por Bagozzi et al. como medida del brand love:

TABLA 4: ÍTEMS DE LA ESCALA DE BRAND LOVE

Factores	Subfactores
<i>Factor de orden superior:</i> Integración de la marca con el yo	<ul style="list-style-type: none"> • Identidad propia actual • Identidad propia deseada • Significado de la vida y recompensas intrínsecas • Fuerza de actitud
<i>Factor de orden superior:</i> Comportamientos impulsados por la pasión	<ul style="list-style-type: none"> • Voluntad de invertir recursos • Deseo pasional de usarlo • Involucración en el pasado
<i>Factor de orden superior:</i> Conexión emocional positiva	<ul style="list-style-type: none"> • Afinidad intuitiva • Apego emocional • Afecto positivo
Relación a largo plazo	
Angustia por separación anticipada	
Valencia actitudinal	

Fuente: Bagozzi et al., 2017

Los autores Bagozzi et al. han decidido subdividir el brand love en seis factores distintivos de los cuales los tres primeros que aparecen en la tabla se identifican como “*factor de orden superior*”, y cuentan con una serie de subfactores lo que permite una medición más detallada e informada.

Bagozzi et al. consiguen esgrimir el primer factor de orden superior, **integración de la marca con el yo** en cuatro subcategorías que hacen referencia a las cuatro primeras enumeradas en la tabla 4: (i) identidad propia actual; (ii) identidad propia deseada; (iii) significado de la vida y recompensas intrínsecas; (iv) fuerza de actitud. Esto implica que

para que se cumpla dicha dimensión nos encontraríamos ante un perfil de consumidor que se sienta representado por la marca y que esta le haga sentir como desea, de sentido a su vida y, además, suela encontrarse pensando en la marca de manera recurrente.

El segundo factor de orden superior se trata de **comportamientos impulsados por la pasión**, en este caso Bagozzi et al. lo divide en tres subfactores: (v) voluntad de invertir recursos, (vi) deseo pasional de usarlo e (vii) involucración en el pasado. Para que se cumpla esta dimensión, el consumidor debe estar dispuesto a invertir tiempo y dinero en la mejora y perfeccionamiento del producto/marca, desear usar el producto y, además, haber interactuado o estar involucrado con la marca en el pasado.

Posteriormente, los autores identificaron como tercer factor de orden superior la **conexión emocional positiva**, constituido por tres subfactores (viii) afinidad intuitiva, (ix) apego emocional y (x) afecto positivo. Este factor requiere que el consumidor sienta que la marca encaja con sus gustos a la perfección, experimente una conexión emocional con ella y considere la marca como divertida y emocionante.

El cuarto factor identificado por Bagozzi et al. se trata de la **relación a largo plazo**, y, para que este factor se cumpla el consumidor debe creer que la marca formará parte de su vida a largo plazo. Como penúltimo factor, los autores hacen referencia a la **angustia por separación anticipada**, donde se corrobora hasta qué punto el consumidor sentiría ansiedad o aprensión si la marca cesase de existir. Finalmente, el sexto y último factor se trata de medir la **valencia actitudinal**, donde los consumidores reflejan sus sentimientos generales hacia la marca en una escala de 7 puntos, donde cuanto más positivos sean estos, mayor será la puntuación asignada.

3. ANÁLISIS EMPÍRICO

3.1. QUE ES TIKTOK Y CUAL ES SU HISTORIA

Es muy probable que en algún momento hayas escuchado hablar de TikTok, ya sea porque lo hayas utilizado personalmente, por que tus hijos o sobrinos lo utilicen o por que algún amigo o familiar te lo haya recomendado. Esta aplicación ha sido muy polémica durante los últimos años, es por ello por lo que resulta interesante conocer su historia y así ayudarnos a comprender en mayor medida las percepciones que mantienen los usuarios respecto a esta marca en cuestión.

La aplicación pasó a reconocerse como uno de los gigantes de las redes sociales en tiempo récord destacando por su perfil adictivo y altos niveles de *engagement*, se lanzó en septiembre de 2016 por la *startup* china *ByteDance* donde se conoce a la marca con el nombre de *Douyin* (D'Souza, 2024). Pero ¿qué es exactamente TikTok? Se trata de una aplicación de vídeos de formato corto, que permite a todos sus usuarios ser creadores y anima a estos a compartir su pasión y expresión creativa a través de vídeos. El 1 de agosto de 2018 se hizo pública su unión con otra de las aplicaciones de vídeos de formato corto de mayor crecimiento en el panorama actual de ese momento, *musical.ly*, manteniendo el nombre de TikTok y creando una nueva experiencia de usuario unificada ofreciendo mayores capacidades para la creación de los videos (TikTok Newsroom, 2018). Citando a D'Souza (2024), fue aquí cuando TikTok aumentó su crecimiento de manera estratosférica, adquiriendo la lista de 200 millones de cuentas con las que contaba *musical.ly* y que se transfirieron a la nueva aplicación.

Gran parte del crecimiento de la aplicación se originó durante los primeros meses de 2020 debido a la pandemia de COVID-19 cuando las personas fueron puestas en cuarentena y recurrieron a sus redes sociales como fuentes de entretenimiento. En 2021 fue la aplicación más descargada del año y alcanzó mil millones de *monthly-active-users* (MAUs)⁴, en 2023 estaba valorada en 66 mil millones de dólares (convirtiéndose en la startup más valiosa del mundo) y a principios de 2024, la aplicación ya se había descargado más de 4.7 mil millones de veces a nivel mundial y contaba con más de 1.5 mil millones de usuarios activos, posicionándose como la quinta red social más popular del mundo (Sheperd, 2024; D'Souza, 2024; TikTok Community, 2021).

⁴ Es una métrica empleada en marketing y que en español se traduce a 'usuarios activos mensuales'.

Sin embargo, a pesar de su notable éxito TikTok ha sido el foco de muchas polémicas en numerosos países. Tal y como apunta D'Souza (2024) la aplicación está prohibida en la India (como otras muchas aplicaciones chinas), debido a las crecientes tensiones entre las dos naciones y en Pakistán. Además, D'Souza menciona países como Bangladesh o Indonesia donde la aplicación fue prohibida brevemente debido a una serie de acusaciones relacionadas con contenido inapropiado, esta prohibición terminó por revocarse cuando la aplicación se comprometió a retirar cualquier tipo de contenido de esta naturaleza.

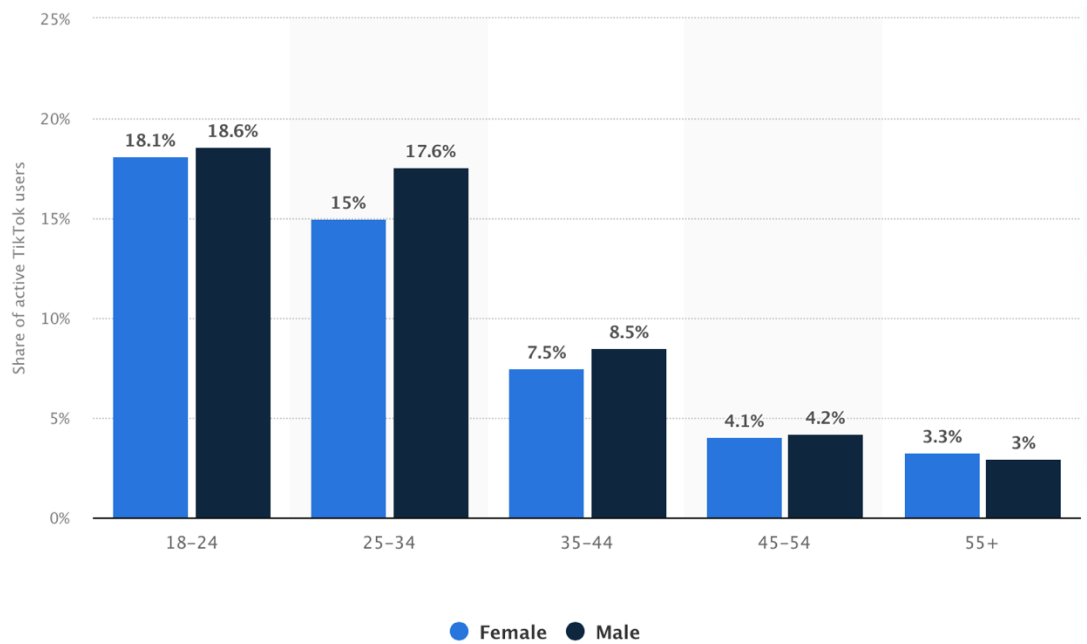
A pesar de impulsar un crecimiento económico significativo para la economía estadounidense en su conjunto, respaldando más de 24.2 mil millones de dólares en producto interno bruto y proporcionar más de 224 mil empleos a estadounidenses en 2023 (TikTok Newsroom, 2024), a lo largo de los años han surgido preocupaciones globales relacionadas con su gestión de la privacidad digital. Durante la presidencia de Trump en 2020, TikTok estuvo cerca de prohibirse debido a preocupaciones sobre la recopilación y uso de los datos de estadounidenses por parte de China. Para evitar restricciones se valoró que ByteDance vendiese el 20% de TikTok a *Oracle* y *Walmart* (empresas estadounidenses) dando lugar a TikTok Global (empresa regulada en EE. UU). En la Unión Europea por otro lado, TikTok se enfrenta a la necesidad de cumplir con la Ley de Servicios Digitales que protege la privacidad de los usuarios, arriesgándose a multas o a su prohibición si esta no se ajusta a sus regulaciones (D'Souza, 2024).

3.2. ANÁLISIS DE LA MARCA

TikTok como cualquier otro servicio cuenta con su propia plataforma de usuarios que se analizará a lo largo de este segmento en más detalle junto con su posicionamiento en el mercado y demás características para obtener una comprensión integral de su dinámica de mercado e imagen de marca y así poder determinar hasta qué punto se trata de una lovebrand o no.

Tal como apuntan los datos recolectados por Ceci (2024), la plataforma ha logrado hacerse hueco sobre todo entre la audiencia más joven, la mayoría de sus usuarios se encuentran concentrados entre los 18 y 24 años y en cuanto a su demográfica, no muestra mucha diferencia entre géneros.

ILUSTRACIÓN 3: DISTRIBUCIÓN DE USUARIOS DE TIKTOK EN TODO EL MUNDO A ENERO DE 2024, POR EDAD Y SEXO



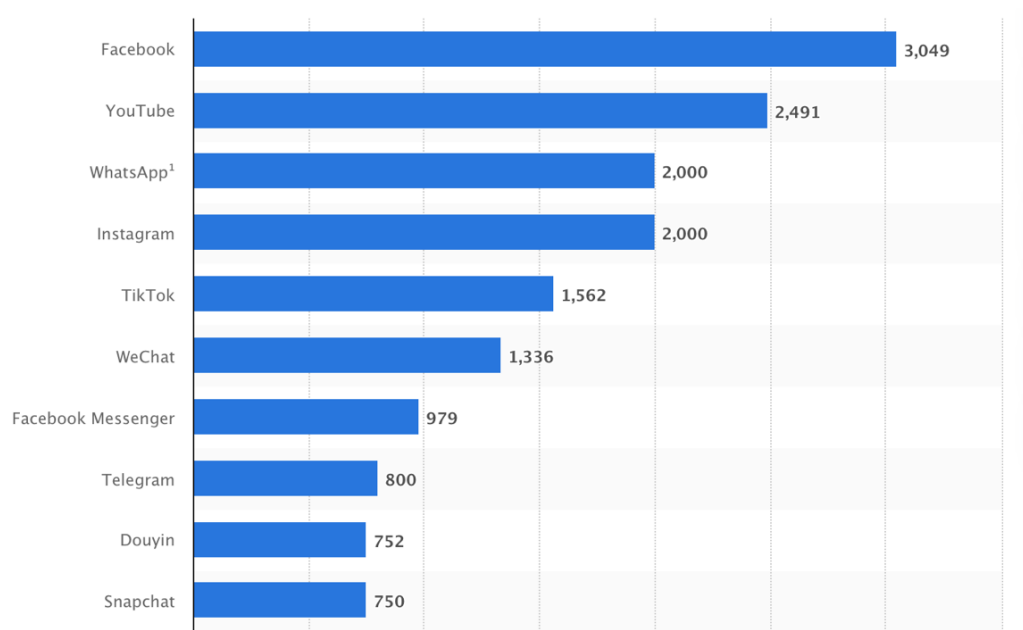
Fuente: Ceci, 2024

De acuerdo con un estudio elaborado por Clear M&C Saatchi donde se analiza una muestra de 4,000 personas en Reino Unido, Francia, Alemania, Italia y España, las cuatro motivaciones principales que impulsan a los usuarios de TikTok a utilizar la plataforma son el **entretenimiento**, la **participación**, la **inspiración** y el **descubrimiento**. De entre estas cuatro, se destaca la de entretenimiento, con un 75% de las personas del estudio asegurando que acuden a la plataforma con el objetivo de sonreír, divertirse y mejorar sus ánimos. Además, el estudio identificó otras tres motivaciones de los usuarios a la hora de usar las plataformas digitales: mantenerse al día, aprender y conectar. Además, cabe mencionar que el 60% de los usuarios confiesan que han experimentado emociones positivas durante el uso de la plataforma, colocando a TikTok entre las dos principales plataformas relacionadas con la positividad (TikTok Noticias, 2021).

Según lo publicado en TikTok Noticias y de acuerdo con el estudio elaborado por Clear M&C Saatchi, se hace referencia a TikTok como un *“hogar para la autenticidad, la positividad y la inclusión”*, donde el 43% de los entrevistados afirman que cualquiera es bienvenido y el 38% se refiere a la plataforma como *“un lugar positivo para estar”*.

Su posición respecto al resto de las redes sociales en el panorama actual y según los datos recabados por Jo Dixon (2024), a principios de 2024, TikTok se posiciona como la quinta red social más popular a nivel mundial en términos de MAUs.

ILUSTRACIÓN 4: REDES SOCIALES MÁS POPULARES A NIVEL MUNDIAL A ENERO DE 2024, CLASIFICADAS POR MAUs (EN MILLONES)



Fuente: Jo Dixon, 2024

Ahora que sabemos cómo TikTok se posiciona frente al resto de sus competidores, cabe mencionar una de las características más valiosas que posee TikTok y que constituye uno de sus secretos para el éxito en esta batalla por la atención. Este activo se trata de su famoso algoritmo de emparejamiento centrado en el usuario y que según Rach & Peter (2021), este supera al de Facebook en captar la atención bajo la teoría del escapismo⁵.

Rach & Peter (2021) destacan que TikTok, con su enfoque en el contenido generado por usuarios (UGC), ha logrado un crecimiento impresionante en términos de MAUs y aumento en el tiempo de uso diario, particularmente entre adolescentes. Los autores también recalcan que el algoritmo (que prioriza la personalización del contenido frente a las conexiones de red) juega un papel fundamental al fomentar el escapismo y permiten a sus usuarios alejarse de la realidad cotidiana fácilmente en comparación con Facebook. Además, Rach & Peter sugieren que la plataforma proporciona un valor de entretenimiento y contenido superior para el usuario individual, lo que podría explicar su éxito en captar y retener a los usuarios por más tiempo.

⁵ Esta teoría se enfoca en la tendencia de los individuos a tratar de distraerse de realidades desagradables a través del entretenimiento o fantasía.

3.3. METODOLOGÍA

3.2.1. Obtención

Una vez analizada la aplicación de TikTok, y su posicionamiento en el mercado, se procederá a realizar un análisis sobre una serie de reseñas realizadas por distintos usuarios que se hayan descargado la aplicación, y que, por tanto, hayan podido disfrutar de los distintos servicios que ofrece la misma. Este enfoque nos permitirá evaluar las distintas emociones y sensaciones que genera la aplicación en los usuarios y de esta misma manera, identificar el valor aportado a los mismos.

En un primer momento se quiso realizar la técnica de *web scrapping*⁶, de manera que se pudiese recabar directamente las opiniones de los distintos usuarios desde la red social X. Sin embargo, actualmente, y debido a una serie de modificaciones recientes en las normativas de privacidad de datos de dicha red, y como consecuencia de la adquisición por parte del empresario Elon Musk, no se ha podido aplicar esta metodología. Por consiguiente, finalmente se optó por recurrir a *Kaggle*, una plataforma web especializada en ciencia de datos, que ofrece, entre otros servicios, la posibilidad de acceder y publicar conjuntos de datos. Más concretamente, los datos se han obtenido de la siguiente página web: [TikTok - Google Play Store Review \(kaggle.com\)](https://www.kaggle.com/datasets/tylertomlinson/tiktok-google-play-reviews)

Para realizar el estudio se ha empleado una base de datos denominada *tiktok_google_play_reviews.csv*, la cual, tras una fase previa de preprocesamiento y limpieza de datos fue convertida a formato Excel e importada a la plataforma RStudio. Allí, se ejecutó una fase definitiva de limpieza de datos antes de pasar a realizar *text mining* de manera que se pudiese obtener aquellas palabras clave y de mayor relevancia sobre las que posteriormente se realizará un análisis de sentimiento.

La base de datos incluye un conglomerado de 406,222 reseñas, cifra que, tras una limpieza de los datos donde se eliminaron los valores nulos, se redujo a 459,657. Estas reseñas fueron publicadas durante el intervalo comprendido entre el 18 de junio de 2022 y el 29 de noviembre de 2022, abarcando así un periodo significativo que permite capturar una amplia gama de opiniones y perspectivas de los usuarios.

⁶ Técnica utilizada para extraer información de sitios web (información personal, comentarios...)

Se seleccionó esta base de datos frente a otras debido a que aportaba un mayor volumen de reseñas, además de recoger opiniones más recientes de manera que el análisis fuese lo más representativo posible de la sociedad actual. Esta selección estratégica garantiza que el estudio se sustente en una muestra representativa y actualizada, proporcionando una visión clara y relevante de las actitudes y opiniones de los usuarios hacia la aplicación.

Cabe mencionar que los datos que contiene la base de datos son comentarios en inglés, por lo que estaríamos analizando las distintas opiniones hacia TikTok dentro de este demográfico específico. Asimismo, se decidió analizar una base de datos que contuviese reseñas en inglés, debido a que solamente el lexicón NRC⁷, junto con algunas excepciones, permite realizar el análisis sobre otros idiomas, de manera que estaríamos limitando el estudio.

3.2.2. Text Mining

Una vez tenemos los datos limpios y preparados para comenzar el análisis nos encontramos con una serie de texto no estructurado, es decir, datos compuestos por una lingüística comprensible para el ser humano, pero no para programas informáticos. Para normalizar estos datos de manera que puedan ser comprendidos por un programa se procede a realizar lo que se conoce como Text Mining. Según definen Nasukawa et al. (2001), el Text Mining consiste en examinar una colección de documentos con el objetivo de descubrir información no contenida en ningún documento individual de la colección; es decir, trata de obtener información sin haber partido de algo.

En este caso, cada reseña supondrá un documento, en total contamos con un número final de 459,657 documentos, los cuales varían en cuanto a número de términos y frases. Al conjunto de estos 459,657 documentos lo denominamos *corpus*, de manera que el *corpus* recogerá los distintos textos sobre los que se realiza el análisis lingüístico. Posteriormente se realiza sobre este corpus un proceso de *tokenización* donde se divide el texto en *tokens* (palabras, en este contexto) y construcción de la matriz *document-term (DTM)*, eliminando *stopwords*⁸, números, puntuación y demás términos irrelevantes, y donde se

⁷ El lexicón NRC es un recurso empleado en el análisis de sentimiento para analizar y descubrir las emociones generadas por un determinado texto.

⁸ Las stopwords son un conjunto de palabras comunes dentro de un idioma específico. Estas palabras no aportan valor informativo al análisis, por lo tanto, resultan redundantes.

representan los documentos en forma de matriz, donde las filas se corresponden con los documentos y las columnas representan cada *token*.

En este estudio, se realiza un proceso de *stemming*, el cual según Khyani et al (2021), produce variantes de una palabra raíz/base. En palabras simples, reduce una palabra base a su palabra de raíz. Utilizamos el *stemming* para acortar la búsqueda y normalizar las oraciones para una mejor comprensión. Una vez realizado el *stemming*, las dimensiones de la matriz se ven afectadas y el número de variables se reduce por 124,753, por lo que resulta útil realizar este proceso en el análisis por varias razones, ya que, se reduce la complejidad del vocabulario del texto, mejora el rendimiento de la búsqueda y ayuda a mantener la uniformidad al tratar variantes de una misma palabra.

3.2.3. Sentiment Analysis

Según un estudio realizado por Cedeno-Moreno y Vargas (2021): *“el análisis de sentimientos es una tarea de procesamiento de lenguaje natural y extracción de información que identifica las opiniones de los usuarios explicadas en forma de comentarios positivos, negativos o neutrales. Varias técnicas pueden ser utilizadas para este fin, por ejemplo, el uso de diccionarios léxicos”*.

En este caso, el análisis se realiza sobre los 16 términos más frecuentes dentro de las reseñas de los distintos usuarios publicadas en Google Play, y para ello, aplicamos tres diccionarios léxicos, Bing, AFINN y NRC. El uso de estos tres lexicones nos proporciona un enfoque comprensivo del análisis de sentimiento. Estos diccionarios se aplican utilizando el paquete *Syuzhet*, este paquete permite la extracción de sentimientos a través de estos lexicones de sentimiento o análisis de sentimiento basado en el aprendizaje automático.

En nuestro caso, en primer lugar, aplicamos el lexicón Bing, desarrollado por Bing Liu y colaboradores, que consiste en una lista de palabras polarizadas como positivas o negativas. No se asigna un grado de sentimiento, sino que se aplica una clasificación binaria de manera que cada palabra en este lexicón se categoriza como positiva o negativa, se trata del lexicón más simple de los tres que se aplican en el análisis.

También se aplicó posteriormente el lexicón AFINN, creado por Finn Årup Nielsen y que consiste en una lista de palabras en inglés donde a cada una de las palabras le corresponde un valor entero que oscila entre el -5 y 5. Este valor refleja la carga sentimental de cada

palabra, donde a las palabras negativas les corresponde una puntuación negativa y viceversa. Ahora bien, las palabras con una mayor puntuación en valor absoluto representan emociones más fuertes. Este lexicón es utilizado habitualmente en el análisis de texto debido a su simplicidad y facilidad de comprensión. Es particularmente útil para análisis de sentimiento que requieran ser rápidos y eficientes, como por ejemplo en el caso de la minería de opiniones en las redes sociales o plataformas online (como en este caso).

Finalmente, se aplicó el lexicón NRC, creado por Saif Mohammad y Peter Turney como parte del proyecto *Emotion Lexicon*, donde presentan un extenso lexicón de asociación palabra-emoción creado mediante un amplio proyecto de anotación en línea (Mohammad & Turney, 2013). Este lexicón clasifica las palabras en 10 dimensiones emocionales, por un lado, las emociones básicas (confianza, miedo, tristeza, disgusto, enfado, alegría, anticipación y sorpresa) donde cada palabra puede estar asociada a más de una, y, por otro lado, los sentimientos binarios (positivos y negativos). Este lexicón es de gran utilidad a la hora de realizar un análisis de sentimiento más matizado y detallado, capturando una gama más amplia de emociones humanas.

Aplicando estos tres métodos, se logra identificar no solo si las reseñas son por lo general positivas o negativas, sino que también entender la riqueza emocional e intensidad detrás de las opiniones de los usuarios sobre TikTok. Esto nos permite evaluar si la propuesta de valor que ofrece la aplicación está alineada con las percepciones y emociones que esta genera en sus usuarios.

3.4. RESULTADOS DEL ANÁLISIS

3.3.1. Descripción de la base de datos obtenida

La base de datos incluye 406,222 reseñas de distintos usuarios de la aplicación TikTok extraídas de la plataforma Google Play, aplicación desde donde los usuarios de Android pueden descargarse distintas aplicaciones móviles. Cabe mencionar que las reseñas fueron publicadas de manera voluntaria, por lo que el análisis se realizó sobre una serie de opiniones aleatorias de determinados usuarios que estaban dispuestos a compartir las mismas públicamente.

En un principio y previo al análisis, se convirtió el documento *CSV*⁹ en un Excel, para poder manipular los datos de manera más sencilla antes de incorporarlos en RStudio. Posteriormente, se realizó una limpieza y preprocesamiento del documento Excel. En primer lugar, los datos se dividieron en columnas, utilizando el signo de puntuación ‘,’ como delimitador, además, se establecieron los distintos formatos correspondientes a cada variable y se eliminaron aquellas variables con información redundante para el análisis, quedándonos solamente con las 4 variables denominadas *username*, *date*, *content* y *score*. De manera que, a RStudio se importó una tabla compuesta por 406,222 observaciones y 4 variables a la que se le pone el nombre de *misreviews*.

TABLA 5: DESCRIPCIÓN DE LAS PRINCIPALES VARIABLES

VARIABLE	DESCRIPCIÓN
USERNAME	Nombre de usuario.
DATE	Fecha de publicación de la reseña.
CONTENT	Comentario que el usuario deja en la plataforma Google Play.
SCORE	Puntuación que el usuario le pone a la aplicación del 1 al 5.

Fuente: Elaboración Propia

Previo a la importación de datos a RStudio se cargaron las distintas librerías que se emplearán para la realización del análisis. Además, se buscó y se comprobó la existencia de valores nulos que posteriormente se eliminaron empleando las siguientes *queries*¹⁰:

```
#Identificamos si existen valores nulos
anyNA(mis_reviews) #en este caso, sí hay NA

#Eliminamos las observaciones que contengan valores nulos
mis_reviews <- na.omit(mis_reviews) #Eliminamos aquellas observaciones que
contengan NA
```

Esto nos dejó con un total de 459,657 reseñas. A continuación, se muestran las primeras 10 observaciones tras la limpieza:

⁹ Un documento CSV es un conjunto de valores separados por comas.

¹⁰ Una query se trata de una solicitud de información formulada en este caso al programa Rstudio con el objetivo de que nos devuelva un resultado específico.

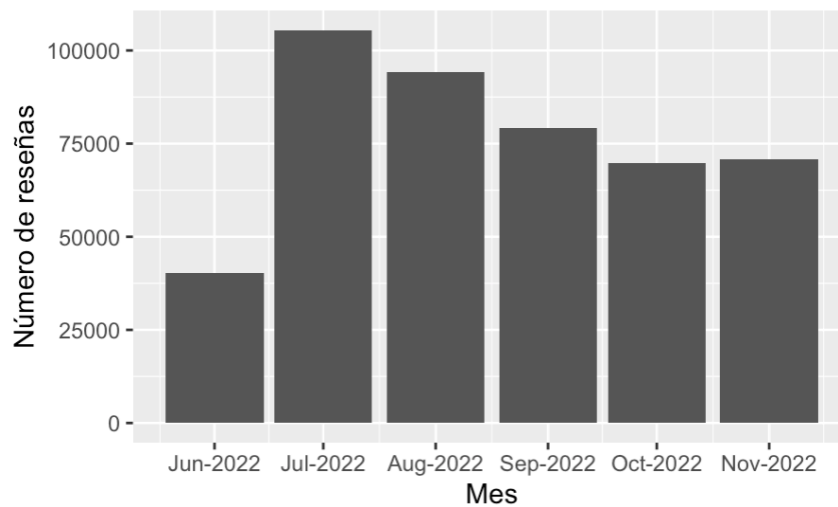
ILUSTRACIÓN 5: IMAGEN DE LAS PRIMERAS DIEZ OBSERVACIONES DE LA MUESTRA

	userName	date	content	score
1	Cassie Moore	2022-11-29	No words	5
2	Kaleb Plummer	2022-11-29	Great fun app so far!	5
3	Rylee Maher	2022-11-29	The app would get a higher rating but I literally can't ...	1
4	Kittykatelyn Romilly	2022-11-29	I WISH I COULD GIVE THIS A 100 PERCENT RATING I L...	5
5	Loveness Malenga	2022-11-29	Pictures and record	5
6	Melvin Crawford	2022-11-29	I love this amazing app	5
7	Muhammad Rehan	2022-11-29	Mohammed Rehan	5
8	Brian Nowak	2022-11-29	Love being on Tik Tok.	5
9	J√/flia Fideles	2022-11-29	Kwai pelo menos da dinheiro	1
10	Rocio Salas	2022-11-29	Post to view is such a step backwards for the platfor...	1

Fuente: Captura de pantalla del análisis realizado en RStudio

En cuanto a las fechas de publicación de las reseñas, como se puede observar en la ilustración 6, estas fueron publicadas durante el intervalo comprendido entre el 18 de junio de 2022 y el 29 de noviembre de 2022.

ILUSTRACIÓN 6: FECHAS DE PUBLICACIÓN DE LA RESEÑAS



Fuente: Captura de pantalla del análisis realizado en RStudio

Antes de comenzar con el análisis se realiza una última limpieza de los datos, de manera que se sustituyen aquellos patrones donde una letra (mayúscula o minúscula) es seguida por un punto y luego inmediatamente otra letra (mayúscula o minúscula), sustituyendo el punto por un espacio. Además, se eliminan también los caracteres especiales y emoticonos, de manera que, se excluye todo excepto caracteres alfanuméricos. Las *queries* que se emplean para ello son las siguientes:

```
#La siguiente función modifica directamente el texto, reemplazando todos los casos
que coincidan con la expresión regular dada.
mis_reviews$content<-gsub("\\.(?=[A-z])", " ",mis_reviews$content, perl = TRUE)

# Quitamos los caracteres especiales y emoticonos. Para eliminar todo excepto
caracteres alfanumericos
mis_reviews$content<-gsub('[^[:alnum:][:blank:]]?&!j\\.|\\-|_|',
"",mis_reviews$content)
```

3.3.2. Text Mining

Tras una limpieza inicial se pasa a la siguiente fase donde cada reseña constituye un documento individual, en primer lugar, creamos un corpus (*micorpus*) donde se recogen todos los documentos (que varían en términos y frases), la fecha, la puntuación correspondiente y una clasificación en función de si la reseña se trata de una opinión positiva o negativa¹¹.

ILUSTRACIÓN 7: IMAGEN DEL CORPUS

```
> summary(micorpus, n=130)
Corpus consisting of 459657 documents, showing 130 documents:
```

Text	Types	Tokens	Sentences	Fecha	Puntuación	Clasificación
text1	2	2	1	2022-11-29	5	POSITIVO
text2	6	6	1	2022-11-29	5	POSITIVO
text3	48	77	4	2022-11-29	1	NEGATIVO
text4	13	17	2	2022-11-29	5	POSITIVO
text5	3	3	1	2022-11-29	5	POSITIVO

Fuente: Captura de pantalla del análisis realizado en RStudio

Como valor informativo en la ilustración 8 se puede observar un recuento de aquellas reseñas clasificadas como negativas y positivas.

ILUSTRACIÓN 8: RECUENTO DE RESEÑAS NEGATIVAS Y POSITIVAS

NEGATIVO	POSITIVO
93529	366128

Fuente: Captura de pantalla del análisis realizado en RStudio

¹¹ Se considera positiva aquella reseña con una puntuación igual a 3 o mayor, y negativa, en caso de que obtuviese una puntuación por debajo del 3.

En este caso el 79,65%¹² de las reseñas se clasificaron como positivas, esto indica una percepción generalmente favorable de la aplicación por parte de los usuarios. Tal porcentaje sugiere que la mayoría de los usuarios están satisfechos con la aplicación, esto nos permite obtener una primera aproximación acerca de la percepción antes de entrar en más detalle.

Una vez tenemos el corpus, realizamos el proceso de tokenización, donde dividimos el texto en palabras y además, eliminamos los números y la puntuación. En la siguiente ilustración podemos observar algunos tokens que se crearon como resultado:

ILUSTRACIÓN 9: TOKENS CREADOS A PARTIR DE LOS PRIMEROS TRES DOCUMENTOS

Name	Type	Value
▼ mistokens	list [459657] (S3: tokens)	List of length 459657
text1	character [2]	'No' 'words'
text2	character [5]	'Great' 'fun' 'app' 'so' 'far'
text3	character [73]	'The' 'app' 'would' 'get' 'a' 'higher' ...

Fuente: Captura de pantalla del análisis realizado en RStudio

Tras tokenizar los documentos, el siguiente paso fue crear la matriz *document-term* (DTM), y que, en este caso inicialmente está compuesta por 459,657 filas y 138,924 *features*. Sin embargo, con el objetivo de agudizar el análisis se eliminaron las stopwords y determinados términos que no aportaban valor al análisis. De manera que las dimensiones de la matriz se vieron reducidas a 138,751 *features*. Las queries empleadas para ello fueron las siguientes:

```
# Quitamos las stopwords de la matriz y vemos los 40 términos más frecuentes
midfm<-dfm_remove(midfm, pattern = stopwords("english"))

# Quitemos además otros términos que no aportan mucho valor en este corpus
midfm<-dfm_remove(midfm, pattern = c("app", "tiktok", "videos", "tik", "tok",
"please", "video", "account", "viral", "just", "can", "dont", "get", "im", "people",
```

¹² Este valor se calculó dividiendo el número de documentos calificados como positivos entre el número total de reseñas.

"viarl", "really", "cant", "time", "much", "apps", "make", "plz", "use", "one", "now", "update", "ever", "team", "even", "üëç", "see", "u", "want", "give", "fix", "go", "got", "help", "really", "views", "phone", "wow", "still", "application", "back", "reason", "know", "id"))

Tras esta limpieza se creó un *data frame* con los 16 términos más frecuentes y su respectiva frecuencia, el cual, se utilizará a la hora de realizar el análisis de sentimiento. Algunas de estas palabras incluyen términos como “good”, “love” o “fun”. El data frame obtenido es el siguiente:

ILUSTRACIÓN 10: IMAGEN DEL DATAFRAME CON LOS DIECISÉIS TÉRMINOS MÁS FRECUENTES Y SU FRECUENCIA

	palabra	frecuencia
1	good	72864
2	nice	46016
3	love	45298
4	like	26462
5	best	22332
6	great	19327
7	fun	14341
8	amazing	12905
9	awesome	7054
10	cool	6752
11	new	6561
12	banned	6206
13	bad	5951
14	problem	5798
15	excellent	4680
16	funny	4324

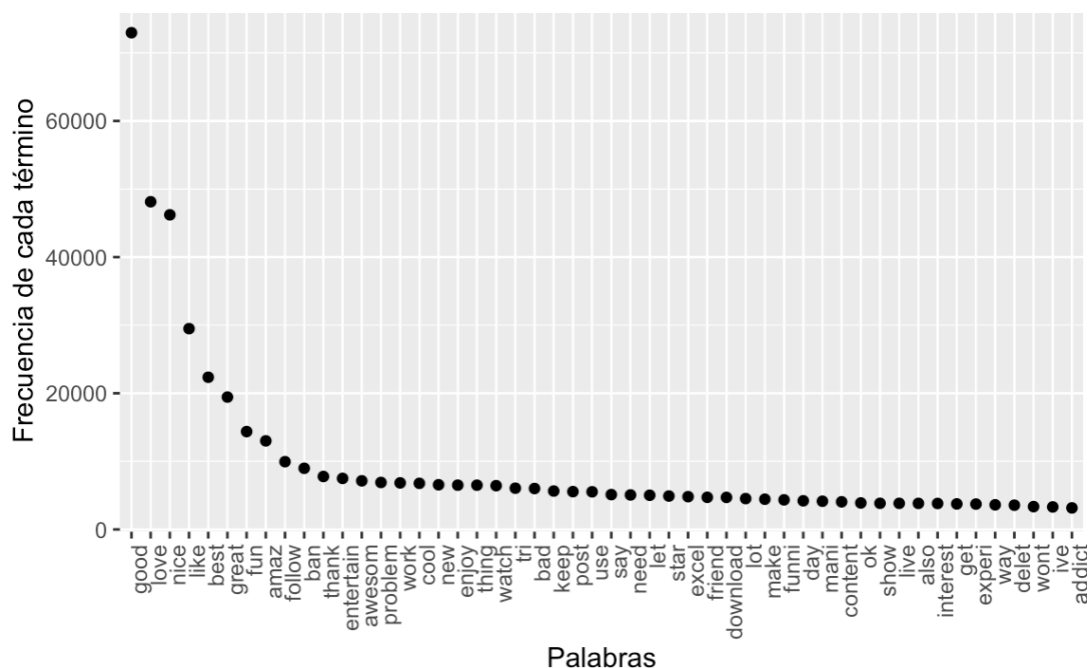
Fuente: Captura de pantalla del análisis realizado en RStudio

Antes de pasar a analizar estos términos en profundidad, se realizó un proceso de stemming, donde los tokens se redujeron a su raíz. En algunos casos este proceso resulta redundante, sin embargo, en lo que concierne a este análisis, el número de features se vio

notablemente reducido por lo que en este estudio sí resulta apropiado. Tras este proceso, se logró un conjunto de datos más manejable y conciso, compuesto por 126,441 features.

Como se menciona en apartados anteriores, una vez realizado el stemming y con los 16 términos más frecuentes, se pasó a realizar un análisis más detallado de los mismos. En la siguiente ilustración se puede observar el eje X, que contiene los 50 términos más frecuentes ordenados de mayor a menor, de izquierda a derecha, según la frecuencia, indicada en el eje Y:

ILUSTRACIÓN 11: GRÁFICO DE LOS CINCUENTA TÉRMINOS MÁS FRECUENTES Y SUS RESPECTIVAS FRECUENCIAS



Fuente: Captura de pantalla del análisis realizado en RStudio

En la gráfica observamos que el término más frecuente es “good”, que se repite un total 72,956 veces en los comentarios. También podemos identificar los términos “love” o “nice” como el segundo y tercer término más repetido, lo que indica que la percepción general de la aplicación es favorable. Además de estos términos positivos, también identificamos términos como “problem” o “bad”, los cuales denotan un sentimiento negativo y que consideramos importante poner en mayor contexto, por lo que más adelante, se analizarán en más detalle algunos de los comentarios que incluyen estos términos con el objetivo de descubrir lo que podrían ser áreas de mejora para Tiktok.

A objeto de profundizar un poco más, se ha creado una nube de palabras con los 100 términos más frecuentes, con el objetivo de obtener *insights* visuales y cualitativos fundamentales para entender la experiencia de los usuarios y la percepción general de TikTok. Los términos en la nube varían en color y tamaño en función de su frecuencia¹³. Para su representación se han indicado dos requisitos, en primer lugar, se ha establecido una frecuencia mínima, de manera que el término aparecerá solamente en caso de que se repita más de 50 veces y una cantidad máxima de palabras, 100, que aparecerán en la nube.

ILUSTRACIÓN 12: GRÁFICO DE NUBE DE PALABRAS



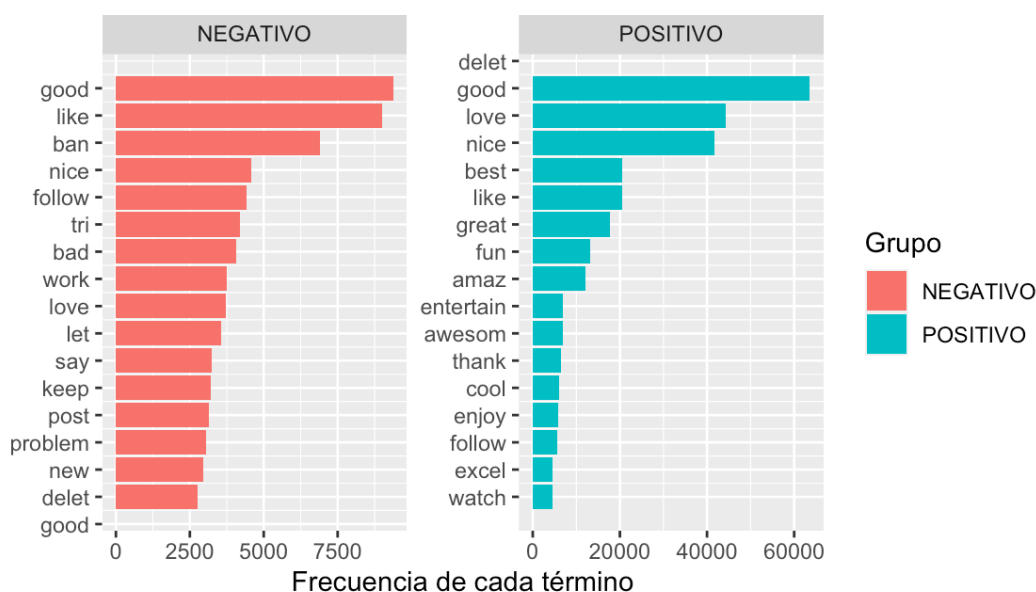
Fuente: Captura de pantalla del análisis realizado en RStudio

La nube de palabras proporciona una primera aproximación cualitativa y representación visual e inmediata de los términos más frecuentes, lo que permite identificar de un vistazo palabras clave y temas predominantes antes de entrar en más detalle, ofreciendo una comprensión rápida de las percepciones de los usuarios y aspectos que más captaron la atención o provocan mayores reacciones.

Sin embargo, es importante para una mayor contextualización, tener en cuenta la frecuencia relativa de los términos positivos frente a los negativos. De esta manera podemos obtener una medida cuantitativa de la percepción de marca. Es por ello por lo que se crearon unos gráficos donde se observan los términos más frecuentes en función de dos grupos, reseñas positivas y reseñas negativas.

¹³ A mayor frecuencia, mayor es el tamaño del término.

ILUSTRACIÓN 13: GRÁFICO DE LOS DIECISÉIS TÉRMINOS MÁS FRECUENTES PARA RESEÑAS POSITIVAS Y NEGATIVAS POR SEPARADO



Fuente: Captura de pantalla del análisis realizado en RStudio

La ilustración 13 nos permite obtener una mejor perspectiva en términos relativos al número de reseñas. En el gráfico izquierdo podemos identificar palabras como “good”, “love” y “like” que normalmente tendrían connotaciones positivas y que aparecen con alta frecuencia en el grupo negativo. Esto podría sugerir que, en el contexto de estas reseñas, estas palabras podrían estar siendo utilizadas con un tono irónico, acompañadas de negaciones o quizás simplemente contengan elementos positivos. Por otro lado, en el gráfico derecho predominan palabras como “good”, “love”, “nice”, “best”, “like”, “great” y “fun”, lo que es coherente con un sentimiento positivo. También aparece el término “delet”, que podría estar derivado de palabras como “delete” o “deleted” (como consecuencia del stemming), su contexto específico requerirá una revisión con el objetivo de contextualizar y entender su inclusión en el grupo de reseñas positivo.

La escala de frecuencias entre los dos grupos es notable, con los términos positivos mostrando una frecuencia mucho mayor, sugiriendo que las expresiones positivas son más comunes en las reseñas publicadas por los usuarios. Es importante matizar que, aunque la visualización nos permite entender las tendencias generales de sentimiento en las reseñas, posteriormente se ha realizado un análisis más detallado del contexto en el

que estos términos se usan, para proporcionar una comprensión aún más matizada de las percepciones de los usuarios hacia TikTok.

A la hora de poner estos términos en contexto para ello se empleará la técnica de análisis *key-words-in-context (KWIC)*¹⁴, esta técnica nos ayuda a entender cómo se utilizan las palabras en el texto, revelando matices de significado que podrían perderse en caso de fijarse solamente en las frecuencias. Un análisis detallado de estos datos nos permitiría entender mejor las inquietudes de los usuarios y a tomar medidas correctivas para mejorar el servicio y políticas de la plataforma. Para ello, se utilizó la función *kwic()* del paquete *quanteda*¹⁵, esta función recibe una palabra y devuelve cada ocurrencia de la misma junto con su contexto inmediato.

En primer lugar, se analizaron las palabras “good”, “love” y “like” dentro del grupo de reseñas negativas para obtener una comprensión más contextualizada. Para ello, se creó un corpus (*corpus_neg*) que incorpora solamente aquellas reseñas que se clasificaron como negativas. La query que se empleó para ello fue la siguiente:

```
# Tanto en las reviews positivas como en las negativas se mencionan las palabras
'good','love' y 'like'.
corpus_neg<-corpus_subset(micorpus, Clasificación=="NEGATIVO")
toks_neg=tokens(corpus_neg)

# ¿Pero qué se dice por ejemplo en las negativas acerca de 'good','love' y 'like'?
kw_neg <- kwic(toks_neg, pattern = c("good","love","like"), window = 12)
head(kw_neg, 30)
```

El resultado fue el siguiente:

ILUSTRACIÓN 14: KEY-WORD-IN-CONTEXT DE LAS PRIMERAS DIECISÉIS RESEÑAS QUE INCLUYEN “GOOD”, “LOVE” Y “LIKE” DENTRO DEL GRUPO DE RESEÑAS NEGATIVAS

¹⁴ Se trata de una técnica empleada en el análisis de sentimiento que devuelve las palabras especificadas y su contexto inmediato.

¹⁵ Herramienta de software del programa RStudio diseñada para emplear en el análisis de texto.


```

rules. Protection of abusive people. It is impossible to maintain | Good |
So I | love | standing in this app with the moderators are not consistent with their
Joined the first video was | like | this app but suddenly when I try to scroll down my saved
| Good | a old scrubber selling herself!!! UNINSTALLED IT!
| Good |
Its so nice and | good | potential too bad for the its community üiÄ
I | like | I really enjoy watching the cool videos
We | like | the tiktok
TF happened? It always shows me stuff that I | like | the app just not the community sexualizing under aged girls making stupid
I dont have trickster voice | like | now its showing me random content 0 interesting for me
so many people and have alot of friends by my videos I | love | all my friends. I need trickster in order to become tiktok
Hey tiktok I do not | like | this app Soo much
I dnt | like | this new design I cant search the tiktoks im looking for
Is this app | good | for sharing
I | love | tik tok you can do whatever you want iv been waiting for

```

Fuente: Captura de pantalla del análisis realizado en RStudio

Tras analizar las 30 primeras reseñas donde aparecen las mencionadas palabras se puede identificar como las palabras “*good*” o “*like*” aparecen en reseñas mostrando opiniones positivas como “*Its so nice and **good** I really enjoy watching the cool videos*”, “*I **like** the tiktok*” o “*I **like** this app its ok ...*”. Sin embargo, también hay situaciones donde las palabras aparecen en comentarios negativos o como problemas expresados por usuarios como “*... I do not **like** this new design...*”, “*...the problem is not solved yetso I will **like** to plead to tiktok for help...*” o “*This app is **good** but one problem I searched the person I cant find...*”. Por otro lado, la palabra “love” también se muestra en diferentes contextos, en su mayoría positivos, como, por ejemplo, “*I **love** tik tok you can do whatever you want...*” o “*I **love** Tiktok*”. Esto sugiere un fuerte afecto o preferencia por la aplicación por parte de los usuarios a pesar de encontrarse en la categoría de reseñas negativas.

Como se mencionó en apartados anteriores, el término “*delet*” aparece con una alta frecuencia en el grupo de reseñas positivas por lo que resulta interesante investigar más de cerca el contexto de estas reviews. Para realizar este análisis se ha asumido que este término viene derivado de la palabra “*delete*” debido al proceso de stemming, por lo que la palabra que insertamos en este caso es esta en lugar de la raíz, debido a que estamos aplicando la función sobre el corpus (*micorpus*). El proceso aplicado fue el mismo, a diferencia de que en este caso se creó un corpus que incluyese solamente las reseñas clasificadas como positivas, una vez se creó se pudo aplicar la técnica sobre el mismo.

Una parte de los resultados se pueden observar a continuación:

ILUSTRACIÓN 15: KEY-WORD-IN-CONTEXT DE LAS PRIMERAS DIECISÉIS RESEÑAS QUE INCLUYEN “DELETE” EN DEL GRUPO DE RESEÑAS POSITIVAS

Los resultados se pueden observar en la siguiente ilustración:

ILUSTRACIÓN 16: KEY-WORD-IN-CONTEXT DE LAS PRIMERAS 16 RESEÑAS QUE INCLUYEN “BANNED” DENTRO DEL GRUPO DE RESEÑAS NEGATIVAS

```

I got | banned | for literally no reason. You guys need to stop
Why do I keep getting | banned | from my brand new account if I havent even done
Dont install it I was | banned | but I did not did anything
It | banned | accounts for no reason and doesnt give you a chance
The app is fun but | banned | your accounts at anytime without warning or opportunity to appeal
Got shadow | banned | for the 3rd timeone stardisrespect tiktok disrespect üöï
very nice app so far although my first account was | banned | its my FIRST TIME using TikTok Im I love it
Destructive app should be | banned | . Terrible takes over phone. Just a bunch of
have copied racists vids as an experiment and got instant | banned | while their vids stayed up. This platform is white
people who are for sure going against guidelines dont get | banned | . Talk about judgment and bullying yeah this app was
even view it you constantly ban accounts that shouldnt be | banned | and then dont get back to the people wanting their
. DO NOT TRUST THIS APP! IT SHOULD BE | BANNED | FOR PRIVACY ISSUES!
app and upload some videos and next day they will | banned | üöö my account please fix this problem
They | banned | me for no reason
I GOT | BANNED | FOR NO REASON
to log back in my account it says im permanently | banned | for multiple guidelines i just want to know whats the

```

Fuente: Captura de pantalla del análisis realizado en RStudio

El resultado muestra que existe un patrón claro de descontento relacionado con las cuentas que han sido prohibidas o suspendidas. Los comentarios reflejan frustración y confusión por parte de los usuarios sobre las políticas de moderación de la plataforma. Los usuarios sienten que las suspensiones han sido injustas o que se han realizado “*for literally no reason*” (literalmente sin ninguna razón) o “*without warning or opportunity to appeal*” (sin advertencia ni oportunidad de apelar). Estos comentarios sugieren que los usuarios consideran que las prohibiciones no se basan en infracciones claras, o directamente, no comprenden las políticas de la plataforma. Esto nos indica que es posible que exista un problema de transparencia y comunicación entre la plataforma y sus usuarios que puede desencadenar en una disminución en el compromiso y posibles daños a la imagen de marca.

A continuación, se profundizó en el análisis a través del uso de n-gramas, Stefanova (2009) define los n-gramas como “*combinaciones de secuencias de categorías gramaticales*”. Más específicamente, se analizaron bigramas, es decir, los pares de palabras consecutivas más frecuentes en las reseñas. De esta manera, se logra una mayor contextualización de las palabras, contexto que podría perderse en caso de analizarse de manera individual. Además, esto nos permitió de la misma manera, desambiguar palabras con múltiples significados como, por ejemplo, “*like*”.

El resultado de los 20 bigramas más frecuentes dentro del conjunto de reseñas se puede observar en la siguiente ilustración:

ILUSTRACIÓN 17: LOS VEINTE BIGRAMAS MÁS FRECUENTES DENTRO DE LAS RESEÑAS Y SUS RESPECTIVAS FRECUENCIAS

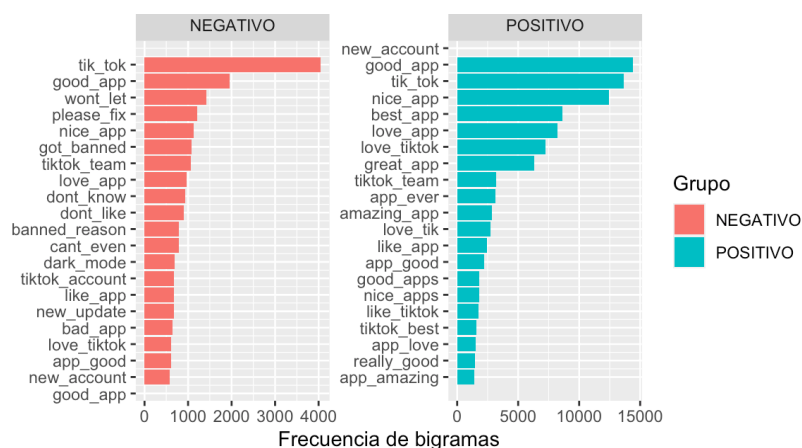
tik_tok	good_app	nice_app	love_app	best_app
17718	16392	13546	9209	9128
love_tiktok	great_app	tiktok_team	app_ever	amazing_app
7872	6848	4241	3607	3128
like_app	love_tik	app_good	like_tiktok	tiktok_account
3127	2928	2811	2111	1982
good_apps	nice_apps	dont_know	really_good	tiktok_best
1958	1946	1766	1706	1688

Fuente: Captura de pantalla del análisis realizado en RStudio

La frecuente presencia de bigramas como por ejemplo “*good_app*” o “*best_app*”, sugiere una percepción generalmente positiva hacia la marca. Esto nos indica que las características o elementos específicos de la marca están siendo valorados y bien recibidos por parte de los usuarios. También resulta importante destacar bigramas como “*love_app*”, debido a que estos indican un sentido de comunidad y señalan un fuerte apego emocional hacia la marca TikTok. Por otro lado, aparece frecuentemente el bigrama “*nice_app*”, el cual sugiere una satisfacción más general.

A pesar de que la gran mayoría de los bigramas más frecuentes son positivos, resulta interesante explorar si existen bigramas con connotaciones negativas que aparezcan con la frecuencia suficiente como para ser de interés, señalando problemas o áreas de mejora. Para tener una aproximación visual y mayor contextualización en lo que concierne a esta observación se realizó el siguiente gráfico:

ILUSTRACIÓN 18: GRÁFICO DE FRECUENCIAS DE BIGRAMAS PARA RESEÑAS POSITIVAS Y NEGATIVAS POR SEPARADO



Fuente: Captura de pantalla del análisis realizado en RStudio

En la ilustración 18 se aprecia una visión general del balance de sentimientos expresados en las reseñas, donde lo primero que salta a la vista, son las frecuencias relativas de los bigramas en cada grupo. Y es que, la frecuencia de bigramas dentro del grupo de reseñas positivas es notablemente superior en comparación con el grupo de reseñas negativas. Este aspecto es de gran relevancia, ya que, las frecuencias relativas de los bigramas en cada grupo pueden proporcionar una aproximación a la medida de intensidad de los sentimientos de los usuarios.

Por un lado, los bigramas del grupo positivo contienen palabras como “*good*”, “*nice*”, “*best*”, “*love*”, con una frecuencia, como se ha mencionado previamente, mucho más alta que en el grupo negativo, lo que indica una experiencia generalmente positiva de los usuarios con la aplicación. Estos bigramas reflejan comentarios positivos generales sobre la aplicación y sugieren una fuerte afición o satisfacción con sus características y funcionalidades.

Por otro lado, los bigramas del grupo negativo incluyen términos como “*wont_get*”, “*please_fix*” o “*dont_like*” que señalan aspectos de insatisfacción por parte de los usuarios o quejas específicas que necesitan ser abordadas, como problemas de usabilidad. Bigramas como “*got_banned*” y “*banned_reason*”, que en conjunto se repiten 1,866 veces en el conjunto de reseñas, demuestran una vez más, que existe un descontento por parte de determinados usuarios por una razón común, y que posiblemente sea un área crítica que requiera atención y que TikTok deba tener en cuenta a la hora de establecer futuras estrategias.

Es importante destacar también que el hecho de que algunos bigramas aparezcan en ambos grupos podría significar que determinados aspectos de la aplicación provocan opiniones divididas entre los usuarios.

3.3.3. Sentiment Analysis

Con el objetivo de profundizar en las emociones y sentimientos que genera la aplicación en sus usuarios se ha realizado un análisis de sentimiento sobre los 16 términos más frecuentes obtenidos en el proceso de Text Mining. Según apuntan Boubeta et al. (2000), “*el verdadero valor de una marca reside en el consumidor, en la importancia que éste le otorga, a partir de los beneficios que la marca le proporciona*”, de manera que explorando los sentimientos que genera la aplicación en los usuarios se nos permite tener

una idea más clara en cuanto al valor que le aporta TikTok a sus consumidores, y de esta misma manera evaluar si las percepciones del público están alineadas con la imagen de marca y valor que la aplicación TikTok quiere transmitir.

Para ello, se aplican 3 tipos de lexicón, en primer lugar, empleamos el lexicón Bing. Este lexicón contiene una lista de palabras positivas y negativas de manera que, la palabra es clasificada de forma binaria en cuanto a su polaridad, positiva (1), neutral (0) o negativa (-1). Los resultados de la asignación fueron los siguientes:

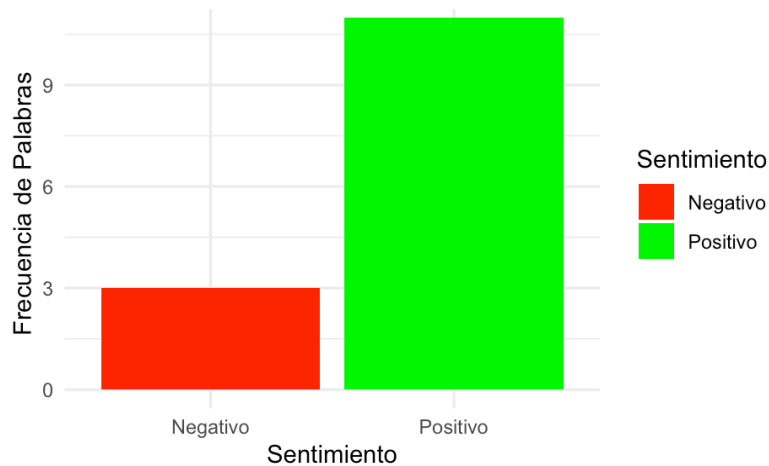
TABLA 6: PUNTUACIONES DE POLARIDAD ASIGNADAS A CADA PALABRA (LEXICÓN BING)

good	1	awesome	1
nice	1	cool	1
love	1	new	0
like	1	banned	0
best	1	bad	-1
great	1	problem	-1
fun	1	excellent	1
amazing	1	funny	-1

Fuente: Elaboración Propia

Este lexicón nos permite una evaluación rápida y clara de los sentimientos de los usuarios sin los matices de intensidad o grado de sentimiento. En cuanto al conteo de palabras positivas y negativas los resultados fueron de 11 y 3 respectivamente, por lo que podemos asegurar que existe una tendencia positiva y que en términos generales la percepción de los usuarios en cuanto a la plataforma es positiva.

ILUSTRACIÓN 19: CONTEO DE PALABRAS POSITIVAS Y NEGATIVAS (LEXICÓN BING)



Fuente: Captura de pantalla del análisis realizado en RStudio

Más tarde se aplicó el lexicón Afinn, que, a diferencia del Bing, contiene una lista de términos en inglés con puntuaciones de polaridad asignadas. Asigna un valor numérico a cada palabra según la intensidad del sentimiento dentro de la escala de -5 (negativo) a 5 (positivo), de esta manera podemos medir el grado de sentimiento vinculado a cada uno de los términos.

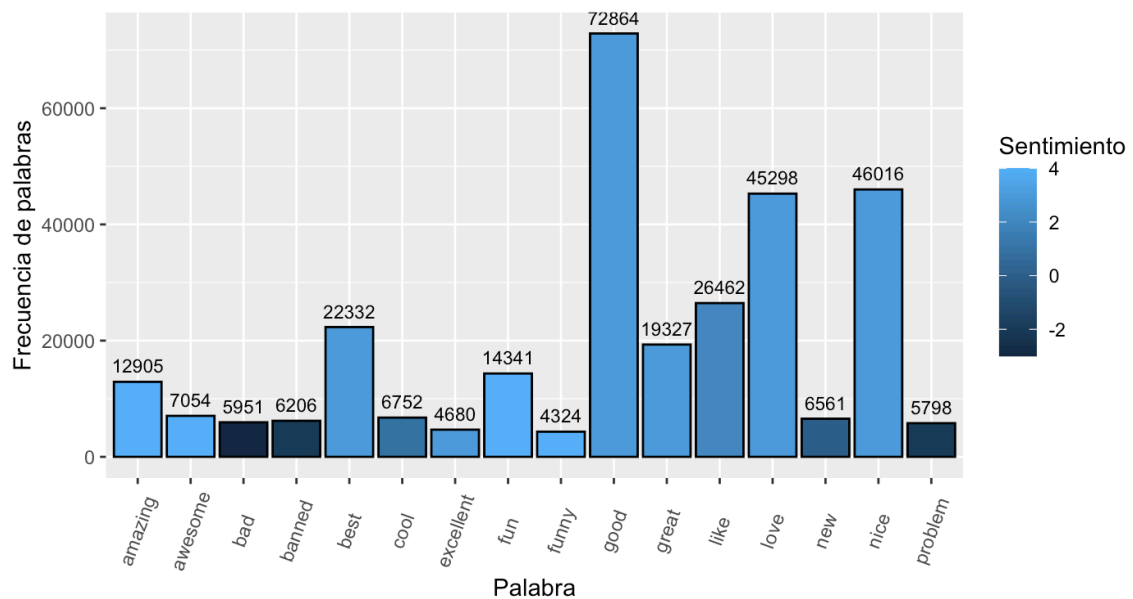
A continuación, se pueden observar las distintas asignaciones asignadas en la tabla 7, junto con su frecuencia en la ilustración 20:

TABLA 7: PUNTUACIONES DE POLARIDAD ASIGNADAS A CADA PALABRA (LEXICÓN AFINN)

good	3	awesome	4
nice	3	cool	1
love	3	new	0
like	2	banned	-2
best	3	bad	-3
great	3	problem	-2
fun	4	excellent	3
amazing	4	funny	4

Fuente: Elaboración Propia

**ILUSTRACIÓN 20: PUNTUACIONES DE POLARIDAD ASIGNADAS A CADA PALABRA
(LEXICÓN AFINN)**



Fuente: Captura de pantalla del análisis realizado en RStudio

En la tabla 7 e ilustración 20 podemos comprobar de nuevo que los términos positivos tienen una presencia más fuerte en el conjunto de datos, ya que, la mayoría de los términos tienen una connotación positiva con calificaciones que varían de 1 a 4, lo que también indica diferentes grados de sentimiento.

Las palabras con la mejor calificación fueron “*awesome*”, “*fun*” y “*funny*”, estas son palabras que suelen asociarse con el humor o disfrute. Además, términos como estos nos indican que los usuarios, no solo están satisfechos con la plataforma, sino que también están emocionalmente comprometidos con ella.

Sin embargo, las palabras “*bad*”, “*banned*” y “*problem*” fueron calificadas con valores de -2 a -3. Estas fueron las únicas que recibieron una calificación negativa, e indican aspectos de una experiencia que ha podido provocar decepción, frustración o molestias en los usuarios. A pesar de que estos términos no son tan frecuentes, en comparación con los anteriormente mencionados, no dejan de ser relevantes para el análisis, ya que, representan áreas críticas que requieren la atención y mejora por parte del equipo de TikTok.

Por último, se aplicó el lexicon NRC, un lexicon más amplio que los anteriores ya que, no solo tiene polaridades positivas o negativas, sino que también asigna polaridades en función de las 8 emociones (enfado, anticipación, disgusto, felicidad, tristeza, sorpresa, confianza y miedo). Los resultados se pueden observar en la siguiente tabla:

TABLA 8: PUNTUACIONES DE POLARIDAD ASIGNADAS A CADA PALABRA SEGÚN LA EMOCIÓN (LEXICÓN NRC)

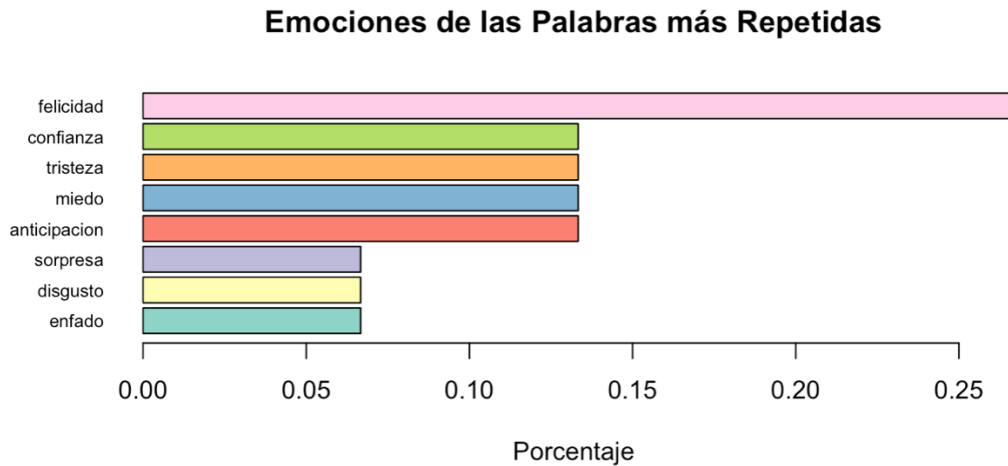
Palabra	Enfado	Anticipación	Disgusto	Felicidad	Tristeza	Sorpresa	Confianza	Miedo
good	0	1	0	1	0	1	1	0
nice	0	0	0	0	0	0	0	0
love	0	0	0	1	0	0	0	0
like	0	0	0	0	0	0	0	0
best	0	0	0	0	0	0	0	0
great	0	0	0	0	0	0	0	0
fun	0	1	0	1	0	0	0	0
amazing	0	0	0	0	0	0	0	0
awesome	0	0	0	0	0	0	0	0
cool	0	0	0	0	0	0	0	0
new	0	0	0	0	0	0	0	0
banned	0	0	0	0	0	0	0	0
bad	1	0	1	0	1	0	0	1
problem	0	0	0	0	1	0	0	1
excellent	0	0	0	1	0	0	1	0
funny	0	0	0	0	0	0	0	0

Fuente: Elaboración Propia

Esto proporciona una visión más detallada de las distintas emociones asociadas con los términos en lugar de simplemente devolver su polaridad general. Resulta muy útil por motivos estratégicos y prácticos, ya que puede ser un punto de partida para TikTok a la

hora de establecer decisiones futuras. En la siguiente figura se pueden observar los porcentajes de palabras asociadas a cada emoción:

ILUSTRACIÓN 21: PORCENTAJES DE LAS PALABRAS ASOCIADAS A LOS OCHO SENTIMIENTOS (LEXICÓN NRC)



Fuente: Captura de pantalla del análisis realizado en RStudio

Según los resultados, parece ser que la emoción que predomina es la de felicidad con un 25%¹⁷ de asociación. La felicidad es una emoción muy fuerte y esta puede mejorar la percepción general de la marca, aumentando el valor y ventaja competitiva de la plataforma frente al resto de competidores. La felicidad de un usuario asociada a la marca suele estar relacionada con el nivel de compromiso que este tiene con la misma, pasando más tiempo en la plataforma y ayudando a construir una reputación sólida y confiable. El hecho de que la emoción con mayor porcentaje sea esta, nos demuestra el afecto que el usuario ha desarrollado por la plataforma.

En segundo lugar, las emociones con una mayor tasa de asociación fueron la confianza, la anticipación, la tristeza y el miedo. Cabe mencionar, que no se está teniendo en cuenta la frecuencia de estos términos a la hora de establecer el porcentaje de palabras asociadas a cada emoción. Por lo que se debe tener en cuenta que palabras como “good”, “love” o “fun” que fueron asociadas a la felicidad, cuentan con una mayor frecuencia que otras palabras como “bad” o “problem” (asociadas a las emociones más negativas).

¹⁷ Este valor se calculó dividiendo el número de términos asociados a la felicidad, 4, entre el total de términos, 16.

4. CONCLUSIONES

En el presente estudio se ha explorado en detalle el fenómeno de *lovebrands*, marcas que trascienden de lo habitual estableciendo conexiones y cultivando una lealtad que va más allá de lo convencional. Una primera parte teórica más el análisis de TikTok como caso de estudio nos han permitido establecer una serie de conclusiones:

- 1) El fenómeno *lovebrand* se trata de un concepto que a lo largo de los años ha ido cobrando cada vez mayor importancia y supone un objetivo para la mayoría de las marcas. Este concepto ha ido evolucionando con el tiempo, en gran parte a raíz de la digitalización por lo que las empresas deben ser lo suficientemente dinámicas como para adaptarse a este entorno de constante cambio si quieren alcanzar esta condición.
- 2) En esta era digital han surgido nuevos modelos de negocio digitales como es el caso de *TikTok*, foco principal de este análisis, que no solo ha adoptado un modelo de negocio digital centrado en el consumidor, sino que también ha sabido crear un espacio donde los propios usuarios pueden tanto consumir contenido como participar activamente en su creación. Esta participación por parte de los consumidores es clave a la hora de desarrollar sentimientos de comunidad, amor y lealtad hacia la marca, factores característicos de las *lovebrands*.
- 3) Las nuevas tecnologías han puesto al alcance de las marcas la posibilidad de obtener una primera aproximación en cuanto a la percepción de su imagen de marca a través de las reseñas online. Este *word-of-mouth* digital puede ser muy útil a la hora de construir amor a la marca, como se ha demostrado en el análisis de reseñas de *TikTok* obtenidas de *Google Play*.
- 4) El empleo de técnicas como el *Text Mining* o *Sentiment Analysis* proporcionan una visión clara de la percepción y emociones que genera la marca en sus usuarios lo que permite a estas marcas actuar y tomar decisiones más informadas. En cuanto al caso TikTok, objeto de estudio de este trabajo los resultados nos permiten llegar a varias conclusiones:
 - a. Entre los términos más frecuentes aparece el término “*love*” (amor), lo que refleja la **pasión por la marca** constituyendo de igual manera una clara **declaración de amor hacia TikTok**, dos de las cinco características del amor a la marca establecidas por Carrol y Ahuvia (2006). Otro término

que surge repetidamente es “*fun*” (divertido) dando a entender que se trata de una **experiencia positiva** para los usuarios, aspecto que resuena con los valores que la marca quiere transmitir a sus usuarios.

- b. El análisis demuestra que se cumplen otras dos de las características de Carroll y Ahuvia (2006) dado que por lo general existe una **evaluación positiva** hacia TikTok por parte de sus usuarios, además de llevar vinculadas **emociones positivas**. Esto se deduce tras el análisis de sentimiento realizado, donde el lexicón Afinn asoció a las palabras más comunes sentimientos fuertes mayoritariamente positivos. Por otro lado, el lexicón NRC demostró que el sentimiento que más predominaba asociado a las palabras fue el de *felicidad*, seguido del sentimiento de *confianza*, ambos sentimientos característicos de una relación de amor.
- c. También se descubrieron áreas de mejora para TikTok en vista de que el lexicón NRC reveló que aproximadamente un 15% de las palabras más frecuentes están asociadas a sentimientos de miedo o tristeza que vienen inferidos de los términos “*bad*” y “*banned*”. Estas reseñas constituyen una corriente de descontentos frente a ciertas políticas de la marca.

Gracias a estos descubrimientos podemos defender que TikTok se encuentra muy cerca de convertirse en una *lovebrand* a nivel global. Los hallazgos revelan que no solo ha captado el interés y la afinidad de sus usuarios a través de herramientas de diferenciación como su algoritmo, sino que también ha fomentado una comunidad vibrante y participativa sobre todo entre los más jóvenes. No obstante, es importante gestionar las percepciones negativas de sus usuarios para consolidar su posición como *lovebrand* asegurando una comunicación transparente y proactiva que fortalezca la lealtad de estos. De esta manera, TikTok podrá no solo mantener, sino aumentar su relevancia en un mercado cada vez más competitivo y saturado, demostrando que las *lovebrands* pueden surgir y prosperar incluso en las circunstancias más desafiantes.

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Sonsoles Martínez González, estudiante de ADE y Business Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "Una lovebrand en la era digital: el caso TikTok", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación [el alumno debe mantener solo aquellas en las que se ha usado ChatGPT o similares y borrar el resto. Si no se ha usado ninguna, borrar todas y escribir “no he usado ninguna”]:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
3. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
4. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
5. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
6. **Generador de problemas de ejemplo:** Para ilustrar conceptos y técnicas.
7. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
8. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: Abril de 2024

Firma: _____

5. BIBLIOGRAFÍA

- Aguado, P. (2016, 8 de noviembre). ¿Cómo convertimos nuestra marca en una "Love Brand"? *Marketing Directo*. Recuperado de <https://www.marketingdirecto.com/especiales/enamorando-al%20consumidor/convertimos-nuestra-marca-una-love-brand>
- Albert, N., Merunka, D., & Valette-Florence, P. (2008). When consumers love their brands: Exploring the concept and its dimensions. *Journal of Business research*, 61(10), 1062-1075.
- American Marketing Association. (1960). *Marketing definitions: A glossary of marketing terms*. Chicago, IL: American Marketing Association.
- American Marketing Association. (2016). Brand. *AMA website for professional marketers*. Recuperado de <https://marketing-dictionary.org/b/brand>
- Bagozzi, R. P., Batra, R., & Ahuvia, A. (2017). Brand love: development and validation of a practical scale. *Marketing Letters*, 28, 1-14.
- Batra, R., Ahuvia, A., & Bagozzi, R. P. (2012). Brand love. *Journal of marketing*, 76(2), 1-16.
- Bergkvist, L., & Bech-Larsen, T. (2010). Two studies of consequences and actionable antecedents of brand love. *Journal of brand management*, 17, 504-518.
- Boubeta, A. R., Mallou, J. V., Tobío, T. B., & Mangin, J. P. L. (2000). El valor de la marca a partir de su relación con el consumidor. *Psicothema*, 247-254.
- Carroll, B. A., & Ahuvia, A. C. (2006). Some antecedents and outcomes of brand love. *Marketing letters*, 17, 79-89.
- Ceci, L. (2024, 1 de febrero). Distribution of TikTok users worldwide as of January 2024, by age and gender. *Statista*. Recuperado el 13 de abril de 2024, de <https://www.statista.com/statistics/1299771/tiktok-global-user-age-distribution/>
- Cedeno-Moreno, D., & Vargas, M. (2021). Machine learning applied to the sentiment analysis. *Technological I+ D Journal*, 16(2), 1-8.

- Coelho, A., Bairrada, C., & Peres, F. (2019). Brand communities' relational outcomes, through brand love. *Journal of Product & Brand Management*, 28(2), 154-165.
- Colmenares, O. (2007). La marca: su definición, sus elementos y su gestión. *Gestiopolis*. Recuperado de <https://www.gestiopolis.com/la-marca-su-definicion-sus-elementos-y-su-gestion/>
- D'Souza, D. (2024, 15 de febrero). TikTok: What It Is, How It Works, and Why It's Popular. *Investopedia*. Recuperado de <https://www.investopedia.com/what-is-tiktok-4588933#citation-25>
- Del Río Lanza, A. B., Vázquez Casielles, R., & Iglesias Argüelles, V. (2002). El valor de marca: perspectivas de análisis y criterios de estimación.
- España. (2001). Ley 17/2001, de 7 de diciembre, de Marcas. Boletín Oficial del Estado, (núm. 294, publicado el 8 de diciembre de 2001). <https://www.boe.es/buscar/act.php?id=BOE-A-2001-23093>
- Félix, M., & Carlos, L. (2002). Data mining: torturando a los datos hasta que confiesen. *Coordinador del programa de Data mining*.
- Gallart Camahort, V., Callarisa Fiol, L. J., & Sánchez, J. (2019). Concepto de marca y sus perspectivas de análisis: una revisión de la literatura. *Redmarka: revista académica de marketing aplicado*, 23(1), 41-56.
- Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of interactive marketing*, 28(2), 149-165.
- Jiménez-Castillo, D., & Sánchez-Fernández, R. (2019). The role of digital influencers in brand recommendation: Examining their impact on engagement, expected value and purchase intention. *International journal of information management*, 49, 366-376.
- Jo Dixon, S. (2024, 2 de febrero). Most popular social networks worldwide as of January 2024, ranked by number of monthly active users (in millions). *Statista*. Recuperado el 13 de abril de 2024, de

<https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

- Khyani, D., Siddhartha, B. S., Niveditha, N. M., & Divya, B. M. (2021). An interpretation of lemmatization and stemming in natural language processing. *Journal of University of Shanghai for Science and Technology*, 22(10), 350-357.c
- Lindenmeyer, L., Seifried, F., & Varga, A. Love is all you need–Defining the concept of love brands.
- Maurya, U. K., & Mishra, P. (2012). What is a brand? A Perspective on Brand Meaning. *European Journal of Business and Management*, 4(3), 122-133.
- Mohammad, S. M., & Turney, P. D. (2013). Nrc emotion lexicon. *National Research Council, Canada*, 2, 234.
- Montalvo-Arroyave, L. F. (2023). Las Lovebrands: Aproximación al concepto, arquitectura y mediciones. *Cuadernos del Centro de Estudios de Diseño y Comunicación*, (188).
- Nasukawa, T., & Nagano, T. (2001). Text analysis and knowledge mining system. *IBM systems journal*, 40(4), 967-984.
- Rach, M., & Peter, M. K. (2021, May). How TikTok’s algorithm beats facebook & co. for attention under the theory of escapism: a network sample analysis of Austrian, German and Swiss users. In *Digital Marketing & eCommerce Conference* (pp. 137-143). Cham: Springer International Publishing.
- Roberts, K. (2004). *Lovemarks: El Futuro más Allá de las Marcas*. Barcelona: Ediciones Urano S.A.
- Sallam, M. A., & Wahid, N. A. (2015). The effects of satisfaction and brand identification on brand love and brand equity outcome: the role of brand loyalty. *European Journal of Business and Social Sciences*, 4(9), 42-55.
- Sarkar, A. (2011). Romancing with a brand: a conceptual analysis of romantic consumer-brand relationship. *Management & Marketing*, 6(1), 79.

- Sheperd, J. (2024, 26 de febrero). 21 Essential TikTok Statistics You Need to Know in 2024. *Social Sheperd*. Recuperado de <https://thesocialshepherd.com/blog/tiktok-statistics>
- Shimp, T. A., & Madden, T. J. (1988). Consumer-object relations: A conceptual framework based analogously on Sternberg's triangular theory of love. *Advances in consumer research*, 15(1).
- Siabato, M. F. F., & Oliva, E. J. D. (2014). Evolución y caracterización de los modelos de Brand Equity. *Suma de negocios*, 5(12), 158-168.
- Stefanova Spassova, M. (2009). *El potencial discriminatorio de las secuencias de categorías gramaticales en la atribución forense de autoría de textos en español*. Universitat Pompeu Fabra.
- Sternberg, R. J. (1986). A triangular theory of love. *Psychological review*, 93(2), 119.
- Swimberghe, K. R., Astakhova, M., & Wooldridge, B. R. (2014). A new dualistic approach to brand passion: Harmonious and obsessive. *Journal of Business Research*, 67(12), 2657-2665.
- TikTok Community. (2021, 27 de septiembre). Thanks a billion! *TikTok*. Recuperado de <https://newsroom.tiktok.com/en-us/1-billion-people-on-tiktok>
- TikTok Newsroom. (2018, 1 de agosto). Musical.ly and TikTok unite to debut new worldwide short-form video platform upgraded app, titled TikTok, now available globally. *TikTok*. Recuperado de <https://newsroom.tiktok.com/en-us/musical-ly-and>
- TikTok Newsroom. (2024, 13 de marzo). Oxford Economics Reports: SMB's use of TikTok Contributed \$24.2 Billion to U.S. Economy in 2023. *TikTok*. Recuperado de <https://newsroom.tiktok.com/en-us/tiktok-economic-impact-report-2024-smb>
- TikTok Noticias. (2021, 11 de mayo). Un nuevo estudio de TikTok revela cómo es la mentalidad de los usuarios europeos. *TikTok*. Recuperado de <https://newsroom.tiktok.com/es-es/estudio-tiktok-revela-mentalidad-usuarios-europeos>

Toro, D. M. (2007). La importancia de la marca como activo de alto impacto. *Revista Colombiana de Marketing*, 6(8), 15-18.

6. ANEXO

#Script analysis text mining & Sentiment Analysis

```
#####
```

```
##### TEXT MINING: TIKTOK REVIEWS #####
```

```
#####
```

```
rm(list = ls())
```

```
options(encoding = "utf8")
```

```
##### CAMBIAR EL DIRECTORIO DE TRABAJO ###
```

```
setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
```

```
getwd()
```

```
##### INSTALAMOS & CARGAMOS PAQUETES #####
```

```
install.packages("stringr")
```

```
install.packages("dplyr")
```

```
install.packages("rvest")
```

```
install.packages("lubridate")
```

```
install.packages("ggplot2")
```

```
install.packages("udpipe")
```

```
install.packages("textdata")
```

```
install.packages("quanteda")
```

```
install.packages("quanteda.textstats")
```

```
install.packages("syuzhet")
```

```
library(stringr) # string manipulation in R
```

```
library(dplyr)
```

```
library(rvest) # para web scraping
```

```
library(lubridate) # para fechas
```

```
library(quanteda)
```

```
library(ggplot2)
```

```
library(quanteda.textstats)
```

```
library(quanteda.textplots)
```

```
library(readxl)
```

```

library(utf8)
library(udpipe)
library(syuzhet)
library(ggwordcloud)
library(igraph)
library(textrank)

#####
##
##### 1. NOS BAJAMOS LOS DATOS Y VEMOS COMO SON: RESEÑAS
TIKTOK #####
#####
##
mis_reviews = read_excel("tiktok_google_play_reviews.xlsx") #reseñas desde junio
2022 hasta noviembre 2022

#####
#####
##### 2. INSPECCIONAMOS UN POCO EL DATA FRAME Y LIMPIAMOS
ALGUNAS COSAS DEL TEXTO #####
#####
#####
View(mis_reviews)
summary(mis_reviews)

#Identificamos si existen valores nulos
anyNA(mis_reviews) #en este caso, sí hay NA

#Eliminamos las observaciones que contengan valores nulos
mis_reviews <- na.omit(mis_reviews) #Eliminamos aquellas observaciones que
contengan NA

##### Inspeccionamos un poco el dataset 'mis_reviews' #####
str(mis_reviews)

```

```

View(mis_reviews)

mis_reviews$username
mis_reviews$date
mis_reviews$content[1:4]# para ver cómo es el texto de las reviews
mis_reviews$score

# Voy a calcular el numero de reseñas por mes y lo grafico
mis_reviews %>%
  dplyr::count(fdate=floor_date(as.Date(date), "month")) %>%
  ggplot(aes(as.Date(fdate), n))+
  geom_bar(stat="identity")+
  scale_x_date(date_labels = "%b-%Y")+
  ggtitle ("Número de reseñas por mes") +
  labs(x = "Mes", y = "Número de reseñas")

#Puede haber texto con el patrón palabra.palabra. Sustituyendo '.' por espacio.
#Con este código identificamos las instancias en el texto donde se da esta situación
grep("[A-z]\\.[A-z]",mis_reviews$content)# estamos utilizando expresiones regulares

#Quitamos ese punto y sustituimos por espacio
#La siguiente función modifica directamente el texto, reemplazando todos los casos que
coincidan con la expresión regular dada.
mis_reviews$content<-gsub("\\.(?=[A-z])", " ",mis_reviews$content, perl = TRUE)

# Quitamos los caracteres especiales y emoticonos. Para eliminar todo excepto
caracteres alfanumericos
mis_reviews$content<-gsub('[^[:alnum:][:blank:];?&!;\\.-]', "",mis_reviews$content)

#####
### 3.CONSTRUCCIÓN DE CORPUS Y MATRIZ ###
#####
micorpus<-corpus(mis_reviews$content)
summary(micorpus)

```

```

# Utilizamos la función docvar() para añadir información a cada documento del corpus
# Por ejemplo, añado las variables fecha y rating al corpus
# Además puedo añadir otra variable (posneg) diciendo si la reseña es positiva (3
estrellas o más) o negativa
docvars(micorpus, c("Fecha", "Puntuación"))<-mis_reviews[,c(2,4)]
docvars(micorpus, "Clasificación")<-ifelse(micorpus$Puntuación<=3, "NEGATIVO",
"POSITIVO")
summary(micorpus, n=130)

# ¿cuántas reviews positivas y negativas hay?
table(micorpus$Clasificación)

# Tokenización y construcción de la matriz document-term
# Quitamos números, puntuación y stopwords
mistokens<-tokens(micorpus,remove_numbers = TRUE, remove_punct = TRUE )
midfm<-dfm(mistokens) # Construcción de la matriz document-term
dim(midfm) # ¿cuántos términos ha detectado?
featnames(midfm)

# Quitamos las stopwords de la matriz y vemos los 40 términos más frecuentes
midfm<-dfm_remove(midfm, pattern = stopwords("english"))
dim(midfm)
topfeatures(midfm,40)

# Quitemos además otros términos que no aportan mucho valor en este corpus
midfm<-dfm_remove(midfm, pattern = c("app", "tiktok", "videos", "tik", "tok",
"please", "video", "account",
"viral", "just", "can", "dont", "get", "im", "people", "viarl",
"really", "cant",
"time", "much", "apps", "make", "plz", "use", "one", "now",
"update", "ever",
"team", "even", "üëç", "see", "u", "want", "give", "fix", "go",
"got", "help", "really",

```

```

"views","phone","wow","still", "application", "back", "reason",
"know", "id"))
dim(midfm)
topfeatures(midfm,40)

term_frecuentes<-topfeatures(midfm,16) # los 16 términos más frecuentes
#convierto los resultados obtenidos en un data frame
terminos_frecuentes<-data.frame(palabra = names(term_frecuentes), frecuencia =
term_frecuentes, row.names = NULL) %>%
  dplyr::arrange(desc(frecuencia))
View(terminos_frecuentes)

#STEMMING -> el nº de variables disminuye por lo que esta función nos es util.
midfm <- dfm_wordstem(midfm, language = "eng")
dim(midfm)
featnames(midfm)

#####
### 4.TABLAS Y GRÁFICOS DE FRECUENCIAS DE PALABRAS ###
#####

# Un gráfico de la frecuencia con la que aparece cada término
# Para ello hacemos uso de la función textstat_frequency()
# necesitamos instalar y cargar quantda.textstat (si no lo hemos hecho antes)
# Veremos los 50 términos más relevantes en todo el corpus
# frecuencia de cada término y el número de documentos en los que aparece

features_freq<-textstat_frequency(midfm, n = 50)
View(features_freq)
ggplot(features_freq, aes(x = reorder(feature,-frequency), y = frequency)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  labs(x = "Palabras", y = "Frecuencia de cada término")

# Hagamos ahora una nube de palabras.

```



```

# Vamos a sacar los 100 términos más frecuentes, que además aparecen 50 veces o más
set.seed(123)
textplot_wordcloud(midfm, min_count=50, min_size=1,
max_words=100,color=c('pink', 'red','orange', 'purple', 'blue','green'))

### Para tener una perspectiva en términos relativos al número de reviews
# hay que tener en cuenta el número de reviews positivas y negativas que teníamos
# Graficamos la frecuencia de las 15 palabras más frecuentes dentro de cada grupo
(reseñas positivas y negativas)
freq_posneg <- textstat_frequency(midfm, n = 16, groups = Clasificación)
ggplot(data = freq_posneg, aes(x = nrow(freq_posneg):1, y = frequency, fill=group)) +
  geom_bar(stat="identity") +
  facet_wrap(~ group, scales = "free") +
  coord_flip() +
  scale_x_continuous(breaks = nrow(freq_posneg):1,
                    labels = freq_posneg$feature) +
  labs(x = NULL, y = "Frecuencia de cada término", fill = "Grupo")
# Había bastantes menos negativas que positivas

#####
##### 5. VAMOS A PONER UN POCO DE CONTEXTO #####
#####
##### 5.1. Uso de Key Words in context - del paquete quanteda/ función kwic() #####
# Tanto en las reviews positivas como en las negativas se mencionan las palabras
'good','love' y 'like'.
corpus_neg<-corpus_subset(micorpus, Clasificación=="NEGATIVO")
toks_neg=tokens(corpus_neg)

# ¿Pero qué se dice por ejemplo en las negativas acerca de 'good','love' y 'like'?
kw_neg <- kwic(toks_neg, pattern = c("good","love","like"), window = 12)
head(kw_neg, 30)

# y qué dicen de 'delete' en las positivas?
corpus_pos<-corpus_subset(micorpus, Clasificación=="POSITIVO")

```

```

toks_pos=tokens(corpus_pos)
kw_pos <- kwic(toks_pos, pattern = "delete", window = 10)
head(kw_pos, 30)

# y qué dicen de 'banned' en las negativas?
kw_neg <- kwic(toks_neg, pattern = "banned", window = 10)
head(kw_neg, 30)

##### 5.2. N-GRAMAS y en concreto BIGRAMAS #####
# Aplicamos esto a las reseñas para ver los bigramas más frecuentes:
bigrams<-quanteda::tokens(micorpus,remove_punct = TRUE)%>%
  tokens_tolower()%>%
  tokens_remove(stopwords("english"))%>%
  tokens_ngrams(n=2) %>% # le decimos que los tokens formados sean bigramas
  dfm()
topFeatures <- topfeatures(bigrams, 20)
topFeatures

# Gráfico de frecuencias de bigramas para reseñas positivas y negativas por separado:
bigram_posneg <- textstat_frequency(bigrams, n = 20, groups = Clasificación)
ggplot(data = bigram_posneg, aes(x = nrow(bigram_posneg):1, y = frequency,
fill=group)) +
  geom_bar(stat="identity") +
  facet_wrap(~ group, scales = "free") +
  coord_flip() +
  scale_x_continuous(breaks = nrow(bigram_posneg):1,
                    labels = bigram_posneg$feature) +
  labs(x = NULL, y = "Frecuencia de bigramas", fill = "Grupo")

#####
##### 6. ANÁLISIS DE SENTIMIENTO #####
#####

##### 6.1 LEXICON BING #####

```

#Contiene una lista de palabras positivas y negativas.Cada palabra está etiquetada con su polaridad (positiva o negativa).

#podemos ver el diccionario de los distintos métodos de análisis de sentimientos así:

```
bing_dictionary <- get_sentiment_dictionary('bing', language = "english")
```

Realiza el análisis de sentimientos

```
sentimiento<-get_sentiment(terminos_frecuentes$palabra, method="bing")
```

sentimiento

Crear un dataframe con el conteo de palabras positivas y negativas

```
conteo_sentimiento <- data.frame(sentimiento = c("Positivo", "Negativo"),
                                cantidad = c(sum(sentimiento == 1 -> "positive"),
                                              sum(sentimiento == -1 -> "negative")))
ggplot(conteo_sentimiento, aes(x = sentimiento, y = cantidad, fill = sentimiento)) +
  geom_bar(stat = "identity") +
  scale_fill_manual(values = c("Positivo" = "green", "Negativo" = "red")) +
  labs(title = "Análisis de Sentimientos con Lexicon BING",
       x = "Sentimiento", y = "Frecuencia de Palabras", fill="Sentimiento") +
  theme_minimal()
```

6.2 LEXICON AFINN

#Contiene un conjunto de 2,477 palabras con puntuaciones de polaridad asignadas.

#Asigna un valor numérico a cada palabra, donde valores positivos indican positividad y valores negativos indican negatividad.

#podemos ver el diccionario de los distintos métodos de análisis de sentimientos así:

```
afinn_dictionary <- get_sentiment_dictionary('afinn', language = "english")
```

#Los distintos métodos son Afinn, NRC y bing. Vamos a usar el método afinn que devuelve un numero de -5 a 5.

```
sentimiento<-get_sentiment(terminos_frecuentes$palabra, method="afinn")
```

sentimiento

Para entenderlo mejor, analizamos el término 7, que tiene el maximo valor de 4.

```
tokns<-get_tokens(terminos_frecuentes$palabra[7])
```

```
get_sentiment(tokns,method="afinn")
```

```
review4<-data.frame(token=tokns,valor=get_sentiment(tokns,method="afinn") )
```

review4

#--- Volvemos ahora con el análisis de sentimiento de nuestras reseñas

```

combinacion <- cbind(terminos_frecuentes,sentimiento)
View(combinacion)
#hago el plot con sentimientos:
ggplot(combinacion, aes(x = palabra , y = frecuencia, fill=sentimiento)) +
  geom_col(colour="black") +
  geom_text(aes(label = frecuencia), colour = "black", size=3, vjust = -0.7) +
  labs(x = "Palabra", y = "Frecuencia de palabras", fill = "Sentimiento") +
  theme(axis.text.x = element_text(angle = 70, vjust = 0.5))

```

6.3 LEXICON NRC

```

#Más amplio en comparación con BING y AFINN, ya que no solo tiene polaridades
positivas y negativas, sino también asociadas con diferentes emociones.
#Incluye categorías de emociones como alegría, sorpresa, tristeza, disgusto, miedo, etc.
#Proporciona una visión más detallada de las emociones asociadas con las palabras en
lugar de simplemente su polaridad general.

```

```

nrc_dictionary <- get_sentiment_dictionary('nrc', language = "english")
sentimiento<-get_sentiment(terminos_frecuentes$palabra, method="nrc")
sentimiento
# capturamos las emociones con NRC
emociones<-get_nrc_sentiment(combinacion$palabra, language="english")
# Voy a incorporar esas emociones al data frame
combinacion<-combinacion %>% mutate (enfado=emociones$anger,
                                     anticipacion=emociones$anticipation,
                                     disgusto=emociones$disgust,
                                     miedo=emociones$fear,
                                     felicidad=emociones$joy,
                                     tristeza=emociones$sadnes,
                                     sorpresa=emociones$surprise,
                                     confianza=emociones$trust,
                                     negativo=emociones$negative,
                                     positivo=emociones$positive
)
View(combinacion)

```

```

# Voy a considerar sólo las 8 emociones (sin tener en cuenta negativos/positivos)

```

```
# qué porcentaje de palabras tienen asociado el sentimiento trust, surprise ..  
library(RColorBrewer)  
barplot(  
  sort(colSums(prop.table(combinacion[,4:11]))),  
  horiz = TRUE,  
  cex.names = 0.7,  
  col = brewer.pal(n = 8,  
    name = "Set3"),  
  las = 1,  
  main = "Emociones de las Palabras más Repetidas", xlab="Porcentaje"  
)
```